

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:02:28 ; Search time 128.615 Seconds  
(without alignments)  
33.078 Million cell updates/sec

Title: US-10-089-452-27  
Perfect score: 51  
Sequence: 1 KASQDVGTSA 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04: \*  
1: geneseqp1980s: \*  
2: geneseqp1990s: \*  
3: geneseqp2000s: \*  
4: geneseqp2001s: \*  
5: geneseqp2002s: \*  
6: geneseqp2003as: \*  
7: geneseqp2003bs: \*  
8: geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	11	2 AAR97313	Aar97313 Humanised
2	51	100.0	11	4 AAB86084	Aab86084 H. pylori
3	51	100.0	11	4 AAB86052	Aab86052 H. pylori
4	51	100.0	11	8 ADM95102	Adm95102 Murine VN
5	51	100.0	11	8 ADM98311	Adm98311 Humanised
6	51	100.0	106	2 AAR97337	Aar97337 Humanised
7	51	100.0	106	2 AAR97335	Aar97335 Murine NE
8	51	100.0	106	4 AAB86108	Aab86108 H. pylori
9	51	100.0	106	4 AAB86070	Aab86070 H. pylori
10	51	100.0	106	8 ADM95097	Adm95097 Murine VN
11	51	100.0	106	8 ADM95100	Adm95100 Humanised
12	51	100.0	106	8 ADM98320	Adm98320 Murine VN
13	51	100.0	106	8 ADM98322	Adm98322 Humanised
14	51	100.0	115	6 ABR42756	Abg76490 Anti-CEA
15	51	100.0	232	6 ABG76490	Abg76490 Light cha
16	51	100.0	232	7 ADE64204	Adg64204 MN14LC pr
17	51	100.0	232	7 ADF60817	Adf60817 hMN-14 li
18	51	100.0	256	6 ABR42757	Abg76490 Anti-CEA
19	51	100.0	257	6 ABR42758	Abg76490 Anti-CEA
20	51	100.0	261	6 ABR42289	Abg76490 Anti-CEA
21	51	100.0	261	6 ABR42754	Abg76490 Anti-CEA
22	51	100.0	261	6 ABR42754	Abg76490 Anti-CEA
23	51	100.0	262	6 ABR42295	Abg76490 Anti-CEA
24	51	100.0	262	6 ABR42293	Abg76490 Anti-CEA
25	51	100.0	262	6 ABR42298	Abg76490 Anti-CEA

26	51	100.0	262	6 ABR42290	Abg76490 Anti-CEA
27	51	100.0	262	6 ABR42291	Abg76490 Anti-CEA
28	51	100.0	391	6 ABR83649	Abg76490 Anti-CEA
29	51	100.0	443	6 ABG76488	Abg76490 Anti-CEA
30	51	100.0	443	6 ABG76488	Abg76490 Anti-CEA
31	48	94.1	11	5 AAU72806	Abg76490 Anti-CEA
32	48	94.1	11	5 ABB07360	Abg76490 Anti-CEA
33	48	94.1	11	6 AAO29874	Abg76490 Anti-CEA
34	48	94.1	11	7 ADJ79792	Abg76490 Anti-CEA
35	48	94.1	108	2 AAY21817	Abg76490 Anti-CEA
36	48	94.1	115	5 ABB07359	Abg76490 Anti-CEA
37	48	94.1	212	5 AAU72815	Abg76490 Anti-CEA
38	48	94.1	212	6 AAO29883	Abg76490 Anti-CEA
39	48	94.1	212	7 ADJ79836	Abg76490 Anti-CEA
40	48	94.1	213	5 AAU72810	Abg76490 Anti-CEA
41	48	94.1	213	5 AAU72819	Abg76490 Anti-CEA
42	48	94.1	213	5 AAU72818	Abg76490 Anti-CEA
43	48	94.1	213	5 AAU72817	Abg76490 Anti-CEA
44	48	94.1	213	5 AAU72816	Abg76490 Anti-CEA
45	48	94.1	213	6 AAO29886	Abg76490 Anti-CEA

ALIGNMENTS

RESULT 1  
AAR97313  
ID AAR97313 standard; peptide; 11 AA.

AC AAR97313;  
XX  
DT 15-OCT-1996 (first entry)  
XX  
DE Humanised monoclonal antibody light chain CDR.  
XX  
KW Monoclonal antibody; humanised; mouse; framework region; FR: CDR;  
KW complementary determining region; anti-carcinoembryonic antigen; CEA;  
KW diagnosis; imaging; therapy; immune response.

OS Mus musculus.  
XX  
PN WO9611013-A1.  
XX

PD 18-APR-1996.  
XX

XX 28-SEP-1995; 95WO-US011964.

XX 05-OCT-1994; 94US-00318157.

XX (IMMU-) IMMUNOMEDICS INC.

XX Hansen HJ, Armour KL;

XX WPI; 1996-209653/21.

XX New humanised anti-CEA monoclonal antibody - having engrafted murine CDRs, used for diagnosis, imaging and therapy of CEA-producing cancers.

XX Claim 5; Page 38; 62pp; English.

XX New humanised monoclonal antibodies (MABs) comprising the complementary determining regions (CDRs) of a parental murine class III anti-carcinoembryonic (CEA) MAB engrafted to the framework regions (FRs) of a heterologous antibody which can be derived from any species including human, retain the anti-CEA binding specificity of the parental murine MAB but are less immunogenic in a human subject than the parental MAB. The humanised antibodies can be used in diagnosis, imaging and therapy of CEA-producing cancers and patients receiving the humanised antibodies and conjugates show improved therapeutic results, decreased immune responses and decreased immune-mediated adverse effects compared to the parent antibody. This sequence is a murine derived CDR region which corresponds to CDR 1 in the light chain of the humanised MAB. See AAR97313-97333

```
SQ Sequence 11 AA;
Query Match 100.0%; Score 51; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSTVA 11
Db 1 KASQDVGTSTVA 11
|||||

RESULT 3
AAB86052
ID AAB86052 standard; peptide; 11 AA.
XX
AC AAB86052;
XX
DT 17-JUL-2001 (first entry)
XX
DE H. pylori catalase derived antibody HP25/6m/1B5 light chain CDR1.
XX
KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;
KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;
KW antibacterial; complementarity determining region.
XX
OS Unidentified.
XX
PN WO200127613-A2.
XX
PD 19-APR-2001.
XX
PF 12-OCT-2000; 2000WO-EP010058.
XX
PR 12-OCT-1999; 99EP-00120351.
PR 16-MAR-2000; 2000EP-00105592.
PR 31-MAR-2000; 2000EP-00107028.
PR 10-MAY-2000; 2000EP-00110110.
XX
PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.
XX
PI Reiter C, Cullmann G, Heppner P, Ringeis A, Mueller H, Haindl E;
XX
DR WPI; 2001-282087/29.
DR N-PSDB; AAF88054.
XX
PT Detecting infections by acid-resistant microorganisms, particularly for
PT diagnosing Helicobacter pylori, comprises an immunoassay on a fecal
PT sample.
XX
PS Claim 21; Page 16; 89pp; German.
XX
CC This invention describes a novel method for detecting, in a mammal,
CC infection by an acid-resistant microorganism (A) which comprises reacting
CC a fecal sample with: (i) a receptor (R) such that a complex is formed
CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part
CC complex is formed with Ag, and the formation of a complex detected. R are
CC specific for an Ag which, after passage through the intestines, at least
CC in some mammals, retains a native (or corresponding) structure against
CC which the mammal produces antibodies (when immunized or infected with
CC (A), or its extracts, lysates or derived proteins (or fragments) or
CC synthetic peptides). The products of the invention have antibacterial
CC activity. The method is used to diagnose infection by Helicobacter,
CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),
CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the
CC progress of treatment. Receptors, particularly antibodies, directed
CC against Ag can be used therapeutically for treatment of infections. The
CC method requires only one R to provide a reasonably secure diagnosis
CC (although use of two R improves sensitivity), so is relatively
CC inexpensive and more easily standardized. Also it is direct, non-
CC invasive, suitable for automation and may indicate the stage of an
CC infection. This sequence represents a complementarity determining region
CC (CDR) from an antibody generated against a Helicobacter pylori antigen
CC (catalase or beta-urease) which is used to illustrate the method of the
CC invention
XX
SQ Sequence 11 AA;
Query Match 100.0%; Score 51; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SQ Sequence 11 AA;
Query Match 100.0%; Score 51; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
 |||||  
 Db 1 KASQDVGTSA 11  
 |||||

RESULT 4  
 ADM95102  
 ID ADM95102 standard; peptide; 11 AA.  
 XX  
 AC ADM95102;  
 XX  
 DT 15-JUL-2004 (first entry)  
 XX  
 DE Murine MN-14 MAb light chain variable region CDR1.  
 XX  
 KW non-medullary thyroid carcinoma; carcinoembryonic antigen; antibody;  
 KW complementarity determining region; pancreatic cancer; breast cancer;  
 KW ovarian cancer.  
 XX  
 OS Mus sp.  
 XX WO2004032962-A1.  
 PN  
 XX  
 PD 22-APR-2004.  
 XX  
 PF 11-OCT-2002; 2002WO-US032307.  
 XX  
 PR 08-OCT-2002; 2002US-0416531P.  
 XX  
 PA (IMMU-) IMMUNOMEDICS INC.  
 XX  
 PI Goldenberg DM, Hansen HJ;  
 XX  
 DR WPI; 2004-364813/34.  
 XX  
 XX Treating non-medullary thyroid carcinoma or medullary thyroid carcinoma  
 PT comprises administering, either concurrently or sequentially, Class III  
 PT anti-CEA (carcinoembryonic antigen) monoclonal antibody or fragment and  
 PT therapeutic agent(s).  
 XX  
 PS Claim 6; SEQ ID NO 20; 88pp; English.  
 XX  
 CC The invention relates to a method of treating a non-medullary thyroid  
 CC carcinoma or treating medullary thyroid carcinoma involves administering  
 CC to a subject, either concurrently or sequentially, a Class III anti-CEA  
 CC (carcinoembryonic antigen) monoclonal antibody or its fragment and at  
 CC least one therapeutic agent. The antibody or its fragment comprises the  
 CC complementarity-determining regions (CDRs) of a murine MN-14 monoclonal  
 CC antibody. (i) is useful for treating medullary thyroid carcinoma or non-  
 CC medullary thyroid carcinoma such as pancreatic cancer, breast cancer,  
 CC ovarian cancer, in a subject. This sequence corresponds to the CDR1 of  
 CC the murine MN-14 antibody light chain variable region and used to  
 CC generate the humanised/chimeric antibody of the invention.  
 XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 51; DB 8; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.003;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
 |||||  
 Db 1 KASQDVGTSA 11  
 |||||

RESULT 5  
 ADM98311  
 ID ADM98311 standard; peptide; 11 AA.  
 XX  
 AC ADM98311;  
 XX  
 DT 15-OCT-1996 (first entry)  
 XX

DT 15-JUL-2004 (first entry)  
 XX Humanised murine MN-14 antibody light chain variable region CDR1 peptide.  
 DE  
 XX carcinoembryonic antigen; CEA; antibody therapy; anti-CEA; humanised;  
 KW murine; mouse; antibody; MN-14; non-medullary thyroid carcinoma;  
 KW medullary thyroid carcinoma; colorectal cancer; gastric cancer;  
 KW hepatocellular carcinoma; breast cancer; lung cancer;  
 KW antibody-dependent cell-mediated cytotoxicity; complement-mediated lysis;  
 KW apoptosis.  
 XX  
 OS Mus sp.  
 OS Synthetic.  
 XX WO2004032857-A2.  
 PN  
 XX  
 PD 22-APR-2004.  
 XX  
 PF 08-OCT-2003; 2003WO-US031801.  
 XX  
 PR 08-OCT-2002; 2002US-0416531P.  
 XX  
 PR 11-OCT-2002; 2002WO-US032307.  
 PR  
 PR 02-MAY-2003; 2003US-0467161P.  
 XX  
 PA (IMMU-) IMMUNOMEDICS INC.  
 XX  
 XX Goldenberg DM, Hansen HJ;  
 PI  
 XX  
 DR WPI; 2004-340797/31.  
 XX  
 XX Treating non-medullary or medullary thyroid carcinoma, involves  
 PT administering to subject, either concurrently or sequentially, anti-  
 PT carcinoembryonic antigen antibody or its fragment and therapeutic agent.  
 XX  
 PS Claim 8; Page 76; 122pp; English.  
 XX  
 CC This invention relates to methods for treating cancers that express  
 CC carcinoembryonic antigens (CEAs). Specifically, it refers to antibody  
 CC therapy and in particular anti-carcinoembryonic antigen (anti-CEA)  
 CC antibodies that can be used following administration of an agent that  
 CC will increase antibody effector functions or following administration of  
 CC an amount of interferon effective to upregulate CEA expression in tumour  
 CC cells. The present invention describes the anti-CEA antibody as a class  
 CC III humanised murine monoclonal antibody that retains specificity for CEA  
 CC binding and is preferably an MN-14 antibody or fragment thereof.  
 CC Accordingly, hMN-14 can be formulated into a cytostatic pharmaceutical  
 CC composition useful for the treatment of non-medullary or medullary  
 CC thyroid carcinoma, colorectal cancer, gastric cancer, hepatocellular  
 CC carcinoma, as well as breast or lung cancer. Furthermore, it can  
 CC stimulate antibody-dependent cell-mediated cytotoxicity, complement-  
 CC mediated lysis and apoptosis as well as inhibit cell cycle progression,  
 CC cell growth and angiogenesis. This peptide sequence is a CDR region of  
 CC the humanised murine MN-14 antibody of the invention.  
 XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 51; DB 8; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.003;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
 |||||  
 Db 1 KASQDVGTSA 11  
 |||||

RESULT 6  
 AAR97337  
 ID AAR97337 standard; protein; 106 AA.  
 XX  
 AC AAR97337;  
 XX  
 DT 15-OCT-1996 (first entry)  
 XX

DE Humanised antibody variable light chain (MN14HuVK).

XX Monoclonal antibody; humanised; mouse; framework region; FR; CDR;

XX complementary determining region; anti-carcinoembryonic antigen; CEA;

KW diagnosis; imaging; therapy; immune response.

XX Mus musculus.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Binding-site 24..34 /label= CDR 1.

FT Binding-site 50..56 /label= CDR 2.

FT Binding-site 89..96 /label= CDR 3.

FT WO9611013-A1.

PN 18-APR-1996.

PD 28-SEP-1995; 95WO-US011964.

XX 05-OCT-1994; 94US-00318157.

XX (IMMU-) IMMUNOMEDICS INC.

PI Hansen HJ, Armour KL;

XX WPI; 1996-209653/21.

DR N-PSDB; AAT29013.

XX New humanised anti-CEA monoclonal antibody - having engrafted murine CDRs, used for diagnosis, imaging and therapy of CEA-producing cancers.

XX Disclosure; Fig 8; 62pp; English.

XX New humanised monoclonal antibodies (MAbs) comprising the complementary determining regions (CDRs) of a parental murine class III anti-carcinoembryonic (CEA) MAb engrafted to the framework regions (FRs) of a heterologous antibody which can be derived from any species including human, retain the anti-CEA binding specificity of the parental murine MAb but are less immunogenic in a human subject than the parental MAb. The humanised antibodies can be used in diagnosis, imaging and therapy of CEA-producing cancers and patients receiving the humanised antibodies and conjugates show improved therapeutic results, decreased immune responses and decreased immune-mediated adverse effects compared to the parent antibody. See AAR97313-97333 for the CDR sequence and framework regions of the humanised MAbs

XX Sequence 106 AA;

Query Match 100.0%; Score 51; DB 2; Length 106;

Best Local Similarity 100.0%; Pred. No. 0.033;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSSVA 11  
| | | | | | | | | |

Db 24 KASQDVGTSSVA 34

RESULT 7

ID AAR97335 standard; protein; 106 AA.

XX AAR97335;

XX 15-OCT-1996 (first entry)

DT Murine NEMO MN-14 variable region light chain.

DE Monoclonal antibody; humanised; mouse; framework region; FR; CDR;

XX complementary determining region; anti-carcinoembryonic antigen; CEA;

KW diagnosis; imaging; therapy; immune response.

XX Mus musculus.

XX Key Location/Qualifiers

FT Binding-site 24..34 /label= CDR 1.

FT Binding-site 50..56 /label= CDR 2.

FT Binding-site 89..96 /label= CDR 3.

FT WO9611013-A1.

PN 18-APR-1996.

PD 28-SEP-1995; 95WO-US011964.

XX 05-OCT-1994; 94US-00318157.

XX (IMMU-) IMMUNOMEDICS INC.

PI Hansen HJ, Armour KL;

XX WPI; 1996-209653/21.

DR N-PSDB; AAT29011.

XX New humanised anti-CEA monoclonal antibody - having engrafted murine CDRs, used for diagnosis, imaging and therapy of CEA-producing cancers.

XX Example 7; Fig 2; 62pp; English.

XX New humanised monoclonal antibodies (MAbs) comprising the complementary determining regions (CDRs) of a parental murine class III anti-carcinoembryonic (CEA) MAb engrafted to the framework regions (FRs) of a heterologous antibody which can be derived from any species including human, retain the anti-CEA binding specificity of the parental murine MAb but are less immunogenic in a human subject than the parental MAb. The humanised antibodies can be used in diagnosis, imaging and therapy of CEA-producing cancers and patients receiving the humanised antibodies and conjugates show improved therapeutic results, decreased immune responses and decreased immune-mediated adverse effects compared to the parent antibody. See AAR97313-97333 for the CDR sequence and framework regions of the humanised MAbs

XX Sequence 106 AA;

Query Match 100.0%; Score 51; DB 2; Length 106;

Best Local Similarity 100.0%; Pred. No. 0.033;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSSVA 11  
| | | | | | | | | |

Db 24 KASQDVGTSSVA 34

RESULT 8

ID AAB86108 standard; protein; 106 AA.

XX AAB86108;

XX 17-JUL-2001 (first entry)

DT H. pylori catalase antibody HP25/2m/1BS V region light chain.

DE Catalase; beta-urease; antibody; antigen; detection; infection; epitope;

XX acid-resistant microorganism; complementarity determining region; CDR;

XX feces; heavy chain; light chain.

XX Unidentified.

XX WO200127612-A2.

```

XX PD 19-APR-2001.
XX PF 12-OCT-2000; 2000WO-EP010057.
XX PR 12-OCT-1999; 99EP-00120351.
XX PR 16-MAR-2000; 2000EP-00105592.
XX PR 31-MAR-2000; 2000EP-00107028.
XX PR 10-MAY-2000; 2000EP-00110110.
XX PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.
XX PI Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;
XX DR WPI; 2001-282086/29.
XX DR N-PSDB; AAF88154.
XX PT Detecting infections by acid-resistant microorganisms, particularly for
XX PT diagnosing Helicobacter pylori, comprises immunochromatographic detection
XX PT of antigen in feces.
XX PS Claim 31; Fig 4; 90pp; German.
XX CC This invention describes a novel method for detecting infection by an
XX CC acid-resistant microorganism (A), in a mammal, using
XX CC immunochromatography. The method is used to diagnose infection by an acid
XX CC -resistant microorganism (A), in a mammal, such as Helicobacter,
XX CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),
XX CC H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,
XX CC inexpensive and non-invasive, and may indicate the stage of infection. A
XX CC test strip used in the method may include a filter to eliminate particles
XX CC present in the sample and only a single receptor provides a reasonably
XX CC secure diagnosis, with specificity and selectivity improved by detecting
XX CC several epitopes (of catalase) or different antigens (catalase and beta-
XX CC urease). The method can be automated. This sequence represents a
XX CC Helicobacter pylori catalase derived antibody V-region light chain
XX CC fragment used to illustrate the method of the invention
XX SQ Sequence 106 AA;

Query Match 100.0%; Score 51; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11
DB 24 KASQDVGTSA 34

RESULT 9
AAB86070
ID AAB86070 standard; protein; 106 AA.
XX AC AAB86070;
XX DT 17-JUL-2001 (first entry)
XX DE H. pylori catalase antibody HP25/6m/1B5 V-region light chain.
XX KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;
XX KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;
XX KW antibacterial; complementarity determining region; V-region.
XX OS Unidentified.
XX XX WO200127613-A2.
XX PN 19-APR-2001.
XX PD 12-OCT-2000; 2000WO-EP010058.
XX PF 12-OCT-1999; 99EP-00120351.
XX PR 16-MAR-2000; 2000EP-00105592.
XX PR 31-MAR-2000; 2000EP-00107028.
XX PR 10-MAY-2000; 2000EP-00110110.
XX PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.
XX PI Reiter C, Cullmann G, Heppner P, Ringels A, Mueller H, Haindl E;
XX DR WPI; 2001-282087/29.
XX DR N-PSDB; AAF88097.
XX PT Detecting infections by acid-resistant microorganisms, particularly for
XX PT diagnosing Helicobacter pylori, comprises an immunoassay on a fecal
XX PT sample.
XX PS Claim 27; Fig 4; 89pp; German.
XX CC This invention describes a novel method for detecting, in a mammal,
XX CC infection by an acid-resistant microorganism (A) which comprises reacting
XX CC a fecal sample with: (i) a receptor (R) such that a complex is formed
XX CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part
XX CC complex is formed with Ag, and the formation of a complex detected. R are
XX CC specific for an Ag which, after passage through the intestines, at least
XX CC in some mammals, retains a native (or corresponding) structure against
XX CC which the mammal produces antibodies (when immunized or infected with
XX CC (A), or its extracts, lysates or derived proteins (or fragments) or
XX CC synthetic peptides). The products of the invention have antibacterial
XX CC activity. The method is used to diagnose infection by Helicobacter,
XX CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),
XX CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the
XX CC progress of treatment. Receptors, particularly antibodies, directed
XX CC against Ag can be used therapeutically for treatment of infections. The
XX CC method requires only one R to provide a reasonably secure diagnosis.
XX CC (although use of two R improves sensitivity), so is relatively
XX CC inexpensive and more easily standardized. Also it is direct, non-
XX CC invasive, suitable for automation and may indicate the stage of an
XX CC infection. This sequence represents a Helicobacter pylori anti-catalase
XX CC derived antibody HP25/6m/1B5 V-region light chain fragment which is
XX CC described in the method of the invention
XX SQ Sequence 106 AA;

Query Match 100.0%; Score 51; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11
DB 24 KASQDVGTSA 34

RESULT 10
ADM95097
ID ADM95097 standard; protein; 106 AA.
XX AC ADM95097;
XX DT 15-JUL-2004 (first entry)
XX DE Murine MN-14 Mab light chain variable region.
XX KW non-medullary thyroid carcinoma; carcinoembryonic antigen; antibody;
XX KW complementarity determining region; pancreatic cancer; breast cancer;
XX KW ovarian cancer.
XX OS Mus sp.
XX XX WO2004032962-A1.
XX PN 22-APR-2004.
XX PD 11-OCT-2002; 2002WO-US032307.
XX PF 08-OCT-2002; 2002US-0416531P.
XX PR

```

XX (IMMU-) IMMUNOMEDICS INC.  
 XX Goldenberg DM, Hansen HJ;  
 PI WPI; 2004-364813/34.  
 DR N-PSDB; ADM95096.  
 XX  
 XX Treating non-medullary thyroid carcinoma or medullary thyroid carcinoma  
 PT comprises administering, either concurrently or sequentially, Class III  
 PT anti-CEA (carcinoembryonic antigen) monoclonal antibody or fragment and  
 PT therapeutic agent(s).  
 XX  
 XX Disclosure; Fig 12A-B; 88pp; English.  
 XX  
 XX The invention relates to a method of treating a non-medullary thyroid  
 CC carcinoma or treating medullary thyroid carcinoma involves administering  
 CC to a subject, either concurrently or sequentially, a Class III anti-CEA  
 CC (carcinoembryonic antigen) monoclonal antibody or its fragment and at  
 CC least one therapeutic agent. The antibody or its fragment comprises the  
 CC complementarity-determining regions (CDRs) of a murine MN-14 monoclonal  
 CC antibody. (I) is useful for treating medullary thyroid carcinoma or non-  
 CC medullary thyroid carcinoma such as pancreatic cancer, breast cancer,  
 CC ovarian cancer, in a subject. This sequence corresponds to the murine MN-  
 CC 14 antibody light chain variable region.  
 XX  
 XX Sequence 106 AA;  
 XX  
 XX Query Match 100.0%; Score 51; DB 8; Length 106;  
 XX Best Local Similarity 100.0%; Pred. No. 0.033;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KASQDVGTSA 11  
 DB 24 KASQDVGTSA 34  
 |||||  
 RESULT 11  
 ADM95100  
 ID ADM95100 standard; protein; 106 AA.  
 AC ADM95100;  
 XX  
 XX 15-JUL-2004 (first entry)  
 DT Humanised murine MN-14 MAb light chain variable region.  
 XX  
 XX non-medullary thyroid carcinoma; carcinoembryonic antigen; antibody;  
 KW complementarity determining region; pancreatic cancer; breast cancer;  
 KW ovarian cancer.  
 XX  
 XX Mus sp.  
 OS Homo sapiens.  
 XX WO2004032962-A1.  
 PN 22-APR-2004.  
 PD 11-OCT-2002; 2002WO-US032307.  
 XX 08-OCT-2002; 2002US-0416531P.  
 XX (IMMU-) IMMUNOMEDICS INC.  
 XX Goldenberg DM, Hansen HJ;  
 PI WPI; 2004-364813/34.  
 DR N-PSDB; ADM95096.  
 XX  
 XX Treating non-medullary thyroid carcinoma or medullary thyroid carcinoma  
 PT comprises administering, either concurrently or sequentially, Class III  
 PT anti-CEA (carcinoembryonic antigen) monoclonal antibody or fragment and  
 PT therapeutic agent(s).

XX Claim 14; Fig 16A; 88pp; English.  
 XX  
 XX The invention relates to a method of treating a non-medullary thyroid  
 CC carcinoma or treating medullary thyroid carcinoma involves administering  
 CC to a subject, either concurrently or sequentially, a Class III anti-CEA  
 CC (carcinoembryonic antigen) monoclonal antibody or its fragment and at  
 CC least one therapeutic agent. The antibody or its fragment comprises the  
 CC complementarity-determining regions (CDRs) of a murine MN-14 monoclonal  
 CC antibody. (I) is useful for treating medullary thyroid carcinoma or non-  
 CC medullary thyroid carcinoma such as pancreatic cancer, breast cancer,  
 CC ovarian cancer, in a subject. This sequence corresponds to the humanised  
 CC murine MN-14 antibody light chain variable region.  
 XX  
 XX Sequence 106 AA;  
 XX  
 XX Query Match 100.0%; Score 51; DB 8; Length 106;  
 XX Best Local Similarity 100.0%; Pred. No. 0.033;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KASQDVGTSA 11  
 DB 24 KASQDVGTSA 34  
 |||||  
 RESULT 12  
 ADM98320  
 ID ADM98320 standard; protein; 106 AA.  
 AC ADM98320;  
 XX  
 XX 15-JUL-2004 (first entry)  
 DT Murine MN-14 antibody light chain variable region protein.  
 XX  
 XX carcinoembryonic antigen; CEA; antibody therapy; anti-CEA; murine; mouse;  
 KW antibody; MN-14; non-medullary thyroid carcinoma;  
 KW medullary thyroid carcinoma; colorectal cancer; gastric cancer;  
 KW hepatocellular carcinoma; breast cancer; lung cancer;  
 KW antibody-dependent cell-mediated cytotoxicity; complement-mediated lysis;  
 KW apoptosis.  
 XX  
 XX Mus sp.  
 OS WO2004032857-A2.  
 PN 22-APR-2004.  
 PD 08-OCT-2003; 2003WO-US031801.  
 XX 08-OCT-2002; 2002US-0416531P.  
 PR 11-OCT-2002; 2002WO-US032307.  
 PR 02-MAY-2003; 2003US-0467161P.  
 XX (IMMU-) IMMUNOMEDICS INC.  
 XX Goldenberg DM, Hansen HJ;  
 PI WPI; 2004-340797/31.  
 DR N-PSDB; ADM98319.  
 XX  
 XX Treating non-medullary or medullary thyroid carcinoma, involves  
 PT administering to subject, either concurrently or sequentially, anti-  
 PT carcinoembryonic antigen antibody or its fragment and therapeutic agent.  
 XX  
 XX Disclosure; Fig 12; 122pp; English.  
 XX  
 XX This invention relates to methods for treating cancers that express  
 CC carcinoembryonic antigens (CEAs). Specifically, it refers to antibody  
 CC therapy and in particular anti-carcinoembryonic antigen (anti-CEA)  
 CC antibodies that can be used following administration of an agent that  
 CC will increase antibody effector functions or following administration of  
 CC an amount of interferon effective to upregulate CEA expression in tumour

CC cells. The present invention describes the anti-CEA antibody as a class  
 CC III humanised murine monoclonal antibody that retains specificity for CEA  
 CC binding and is preferably an MN-14 antibody or fragment thereof.  
 CC Accordingly, hMN-14 can be formulated into a cytostatic pharmaceutical  
 CC composition useful for the treatment of non-medullary or medullary  
 CC thyroid carcinoma, colorectal cancer, gastric cancer, hepatocellular  
 CC carcinoma, as well as breast or lung cancer. Furthermore, it can  
 CC stimulate antibody-dependent cell-mediated cytotoxicity, complement-  
 CC mediated lysis and apoptosis as well as inhibit cell cycle progression,  
 CC cell growth and angiogenesis. This polypeptide sequence is murine MN-14  
 CC antibody light chain variable region protein of the invention.

XX SQ Sequence 106 AA;

Query Match 100.0%; Score 51; DB 8; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 0.033;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
 |||||  
 Db 24 KASQDVGTSA 34

RESULT 13  
 ADM98322  
 ID ADM98322 standard; protein; 106 AA.

XX AC ADM98322;

XX DT 15-JUL-2004 (first entry)

XX DE Humanised murine MN-14 antibody light chain variable region protein.

XX KW carcinoembryonic antigen; CEA; antibody therapy; anti-CEA; murine; mouse;  
 KW antibody; MN-14; non-medullary thyroid carcinoma;  
 KW medullary thyroid carcinoma; colorectal cancer; gastric cancer;  
 KW hepatocellular carcinoma; breast cancer; lung cancer;  
 KW antibody-dependent cell-mediated cytotoxicity; complement-mediated lysis;  
 KW apoptosis; humanised.

XX Mus sp.

OS Synthetic.

XX WO2004032857-A2.

XX PD 22-APR-2004.

XX PF 08-OCT-2003; 2003WO-US031801.

XX PR 08-OCT-2002; 2002US-0416531P.

XX PR 11-OCT-2002; 2002WO-US032307.

XX PR 02-MAY-2003; 2003US-0467161P.

XX PA (IMMU-) IMMUNOMEDICS INC.

XX Goldenberg DM, Hansen HJ;

XX WPI; 2004-340797/31.

DR N-PSDB; ADM98321.

XX Treating non-medullary or medullary thyroid carcinoma, involves  
 PT administering to subject, either concurrently or sequentially, anti-  
 PT carcinoembryonic antigen antibody or its fragment and therapeutic agent.

XX Disclosure; Fig 23a; 122pp; English.

XX This invention relates to methods for treating cancers that express  
 CC carcinoembryonic antigens (CEAs). Specifically, it refers to antibody  
 CC therapy and in particular anti-carcinoembryonic antigen (anti-CEA)  
 CC antibodies that can be used following administration of an agent that  
 CC will increase antibody effector functions or following administration of  
 CC an amount of interferon effective to upregulate CEA expression in tumour  
 CC cells. The present invention describes the anti-CEA antibody as a class

CC III humanised murine monoclonal antibody that retains specificity for CEA  
 CC binding and is preferably an MN-14 antibody or fragment thereof.  
 CC Accordingly, hMN-14 can be formulated into a cytostatic pharmaceutical  
 CC composition useful for the treatment of non-medullary or medullary  
 CC thyroid carcinoma, colorectal cancer, gastric cancer, hepatocellular  
 CC carcinoma, as well as breast or lung cancer. Furthermore, it can  
 CC stimulate antibody-dependent cell-mediated cytotoxicity, complement-  
 CC mediated lysis and apoptosis as well as inhibit cell cycle progression,  
 CC cell growth and angiogenesis. This polypeptide sequence is the humanised  
 CC murine MN-14 antibody light chain variable region protein of the  
 CC invention.

XX SQ Sequence 106 AA;

Query Match 100.0%; Score 51; DB 9; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 0.033;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
 |||||  
 Db 24 KASQDVGTSA 34

RESULT 14

ABR42756

ID ABR42756 standard; protein; 115 AA.

XX AC ABR42756;

XX DT 26-AUG-2003 (first entry)

XX DE Anti-CEA humanized monoclonal antibody hMN-14 V kappa region.

XX KW Humanized; monoclonal antibody; antibody; hMN-14;  
 KW carcinoembryonic antigen; CEA; tumour; cytostatic; diabody.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO2003033654-A2.

XX PD 24-APR-2003.

XX PF 15-OCT-2002; 2002WO-US032718.

XX PR 15-OCT-2001; 2001US-0328835P.

XX PR 21-DEC-2001; 2001US-0341881P.

XX PR 08-JAN-2002; 2002US-0345641P.

XX PR 22-AUG-2002; 2002US-0404919P.

XX PA (ROSS/) ROSSI E.

XX PI Rossi E, Chang CK, Goldenberg DM;

XX WPI; 2003-513460/48.

XX New multivalent, monospecific binding protein comprising two or more  
 PT binding sites having affinity for the same single target antigen, where  
 PT each binding site is associated with scFv fragments, useful for  
 PT diagnosing or treating tumor.

XX Disclosure; Fig 12; 62pp; English.

XX The present sequence is the protein sequence of the light chain variable  
 CC region of anti-carcinoembryonic antigen (CEA) humanized monoclonal  
 CC antibody hMN-14. scFv constructs comprising the hMN-14 V kappa and VH  
 CC regions are used in novel multivalent, monospecific binding proteins  
 CC (diabody, triabody and tetrabody) of the invention (see also ABR42754,  
 CC ABR42757 and ABR42758). Claimed binding proteins are specific for a  
 CC tumour-associated antigen, and are used in the diagnosis and treatment of  
 CC a tumour, carcinoma, melanoma, sarcoma, neuroblastoma, leukaemia, glioma,  
 CC lymphoma or myeloma, or a cancer selected from acute lymphoblastic  
 CC leukaemia, acute myelogenous leukaemia, biliary, breast, cervical,

CC chronic lymphocytic leukaemia, chronic myelogenous leukaemia, colorectal,  
CC endometrial, oesophageal, gastric, head and neck, Hodgkin's lymphoma,  
CC lung, medullary thyroid, non-Hodgkin's lymphoma, ovarian, pancreatic,  
CC prostate and bladder  
XX  
SQ Sequence 115 AA;  
Query Match 100.0%; Score 51; DB 6; Length 115;  
Best Local Similarity 100.0%; Pred. No. 0.037; 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KASQDVGTSSVA 11  
Db 24 KASQDVGTSSVA 34  
|||||  
RESULT 15  
ABG76490  
ID ABG76490 standard; protein; 232 AA.  
XX  
AC ABG76490;  
DT 13-MAY-2003 (first entry)  
XX  
DE Light chain variable, VL, region of antibody hMN14.  
XX  
KW CEA; carcinoembryonic antigen; IgTCR; T-cell receptor; cancer; tumour;  
KW colorectal cancer; breast cancer; lung cancer; hMN14; cytostatic; mouse;  
KW human; zeta signalling chain; CD8alpha hinge; humanised antibody;  
KW light chain variable region; VL.  
XX  
OS Homo sapiens.  
OS Mus sp.  
OS Synthetic.  
OS Chimeric.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..19  
FT Protein /label= Signal\_peptide  
FT 20..232  
FT /label= Mature\_VL\_of\_hMN14  
XX  
PN US2002165360-A1.  
XX  
PD 07-NOV-2002.  
XX  
PF 10-DEC-2001; 2001US-00006771.  
XX  
PR 30-NOV-2000; 2000US-0250087P.  
PR 30-NOV-2000; 2000US-0250090P.  
XX  
PA (JUNG/) JUNGHANS R P.  
XX  
PI Junghans RP;  
XX  
XX WPI; 2003-298705/29.  
DR N-PSDB; ABX13170.  
XX  
PT New chimeric molecule from humanized antibody against carcinoembryonic  
PT antigen and having signaling molecules of T cells and other effector  
PT cells, useful for the treatment of colorectal, breast and lung cancers.  
XX  
PS Disclosure; Page 10-11; 20pp; English.  
XX  
CC The invention relates to a chimaeric molecule comprising the  
CC carcinoembryonic antigen (CEA) binding domain of humanised antibody MN14  
CC as a single chain antibody with a (GGSGS)3 linker, the zeta signaling  
CC chain of the T cell receptor (TCR) and an intervening CD8alpha hinge in  
CC which the cysteine residues have been mutated, with the IgTCR molecule  
CC occupying nucleotides 2426-3766 of the retroviral vector sequence.  
CC appearing as ABX13168. The new chimaeric molecule expressed in T cells,  
CC NK (not defined) or other effector cells are useful in treating patients  
CC with cancers expressing the CEA antigen, together with other or with

CC heterologous constructs to engage additional stimulatory and functional  
CC properties of the effector cells to enhance the anti-tumour therapeutic  
CC efficacy. The cancer disorder includes colorectal, breast and lung  
CC cancers. The present sequence represents the light chain variable, VL,  
CC region of antibody hMN14. The VL protein is humanised (complementarity  
CC determining region-grafted) and has a leader peptide  
XX  
SQ Sequence 232 AA;  
Query Match 100.0%; Score 51; DB 6; Length 232;  
Best Local Similarity 100.0%; Pred. No. 0.077; 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KASQDVGTSSVA 11  
Db 43 KASQDVGTSSVA 53  
|||||  
Search completed: August 9, 2005, 14:14:48  
Job time : 129.615 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:07:04 ; Search time 33 Seconds  
(without alignments)  
24.883 Million cell updates/sec

Title: US-10-089-452-27  
Perfect score: 51  
Sequence: 1 KASQDVGTSA 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp:\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp:\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp:\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pcp:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	11	2	US-08-318-157B-20
2	51	100.0	11	4	US-09-253-794-20
3	51	100.0	106	2	US-08-318-157B-4
4	51	100.0	106	2	US-08-318-157B-19
5	51	100.0	106	4	US-09-253-794-4
6	51	100.0	106	4	US-09-253-794-19
7	44	86.3	107	3	US-08-838-682-19
8	44	86.3	107	3	US-08-895-914-19
9	44	86.3	107	3	US-09-357-710A-19
10	44	86.3	107	4	US-09-357-707-19
11	44	86.3	107	4	US-09-357-708-19
12	44	86.3	249	2	US-08-797-689-18
13	44	86.3	249	4	US-09-984-186-18
14	43	84.3	11	2	US-08-602-725-6
15	43	84.3	11	3	US-08-646-265A-118
16	43	84.3	11	3	US-09-406-535-8
17	43	84.3	11	4	US-09-508-413A-8
18	43	84.3	104	4	US-09-508-413A-16
19	43	84.3	107	3	US-08-767-128-32
20	43	84.3	107	3	US-08-646-265A-130
21	43	84.3	107	3	US-08-646-265A-131
22	43	84.3	107	3	US-09-406-535-4
23	43	84.3	109	2	US-08-602-725-30
24	43	84.3	110	2	US-08-602-725-2
25	43	84.3	123	1	US-08-398-613A-20
26	43	84.3	123	1	US-08-398-612A-20
27	43	84.3	123	1	US-08-398-611A-20

28	43	84.3	123	2	US-08-491-334A-20	Sequence 20, Appl
29	43	84.3	123	3	US-09-027-449-17	Sequence 17, Appl
30	43	84.3	123	3	US-08-804-444A-17	Sequence 17, Appl
31	43	84.3	123	3	US-09-026-985-17	Sequence 17, Appl
32	43	84.3	123	4	US-09-121-952A-17	Sequence 17, Appl
33	43	84.3	123	4	US-09-234-340A-17	Sequence 17, Appl
34	43	84.3	126	3	US-08-646-265A-43	Sequence 43, Appl
35	43	84.3	126	3	US-08-646-265A-47	Sequence 47, Appl
36	43	84.3	126	3	US-08-646-265A-51	Sequence 51, Appl
37	43	84.3	126	3	US-08-646-265A-53	Sequence 53, Appl
38	43	84.3	126	3	US-08-646-265A-57	Sequence 57, Appl
39	43	84.3	126	3	US-08-646-265A-59	Sequence 59, Appl
40	43	84.3	126	3	US-08-646-265A-63	Sequence 63, Appl
41	43	84.3	126	3	US-08-646-265A-65	Sequence 65, Appl
42	43	84.3	126	3	US-08-646-265A-69	Sequence 69, Appl
43	43	84.3	126	3	US-08-646-265A-73	Sequence 73, Appl
44	43	84.3	126	3	US-08-646-265A-75	Sequence 75, Appl
45	43	84.3	126	3	US-08-646-265A-77	Sequence 77, Appl

ALIGNMENTS

RESULT 1  
US-08-318-157B-20  
; Sequence 20, Application US/08318157B  
; Patent No. 5874540  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; APPLICANT: ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Poley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/318,157B  
; FILING DATE: 05-OCT-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665  
; REFERENCE/DOCKET NUMBER: 18733/464  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-318-157B-20

Query Match 100.0%; Score 51; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00072;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KASQDVGTSA 11  
Db 1 KASQDVGTSA 11

```
RESULT 2
US-09-253-794-20
; Sequence 20, Application US/09253794
; Patent No. 6676924
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/253,794
; FILING DATE: 22-Feb-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,157
; FILING DATE: 05-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-253-794-20
Query Match 100.0%; Score 51; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00072;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11
Db 1 KASQDVGTSA 11

RESULT 3
US-08-318-157B-4
; Sequence 4, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
```

```
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-318-157B-4
Query Match 100.0%; Score 51; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11
Db 24 KASQDVGTSA 34

RESULT 4
US-08-318-157B-19
; Sequence 19, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
```



;  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-318-157B-19

Query Match 100.0%; Score 51; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 0.0084;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11  
| | | | | | | | | |  
Db 24 KASQDVGTSA 34

## RESULT 5

US-09-253-794-4  
; Sequence 4, Application US/09253794  
; Patent No. 6676924  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/253,794  
; FILING DATE: 22-Feb-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/318,157  
; FILING DATE: 05-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665  
; REFERENCE/DOCKET NUMBER: 18733/464  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 106 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Query Match 100.0%; Score 51; DB 4; Length 106;  
Best Local Similarity 100.0%; Pred. No. 0.0084;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11  
| | | | | | | | | |  
Db 24 KASQDVGTSA 34

## RESULT 6

US-09-253-794-19  
; Sequence 19, Application US/09253794  
; Patent No. 6676924

;  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/253,794  
; FILING DATE: 22-Feb-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/318,157  
; FILING DATE: 05-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665  
; REFERENCE/DOCKET NUMBER: 18733/464  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 106 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-253-794-19

Query Match 100.0%; Score 51; DB 4; Length 106;  
Best Local Similarity 100.0%; Pred. No. 0.0084;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11  
| | | | | | | | | |  
Db 24 KASQDVGTSA 34

## RESULT 7

US-08-838-682-19  
; Sequence 19, Application US/08838682  
; Patent No. 6107090  
; GENERAL INFORMATION:  
; APPLICANT: Bander M.D., Neil H.  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603-1051  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

```
/ APPLICATION NUMBER: US/08/838,682
/ FILING DATE:
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/016,976
/ FILING DATE: 06-MAY-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/022,125
/ FILING DATE: 18-JUL-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Goldman, Michael L.
/ REGISTRATION NUMBER: 30,727
/ REFERENCE/DOCKET NUMBER: 19603/1172
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (716) 263-1304
/ TELEFAX: (716) 263-1600
/ INFORMATION FOR SEQ ID NO: 19:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 107 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-838-682-19

Query Match      86.3%; Score 44; DB 3; Length 107;
Best Local Similarity 90.0%; Pred. No. 0.21;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 KASQDVGTSTV 10
Db      24 KASQDVGTAV 33

RESULT 8
US-08-895-914-19
/ Sequence 19, Application US/08895914
/ Patent No. 6136311
/ GENERAL INFORMATION:
/ APPLICANT: Bander, Neil H.
/ TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CANCER
/ NUMBER OF SEQUENCES: 19
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
/ STREET: Clinton Square, P.O. Box 1051
/ CITY: Rochester
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 14603-1051
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/895,914
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/016,976
/ FILING DATE: 06-MAY-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/022,125
/ FILING DATE: 18-JUL-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/838,682
/ FILING DATE: 09-APR-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Goldman, Michael L.
/ REGISTRATION NUMBER: 30,727
/ REFERENCE/DOCKET NUMBER: 19603/1173
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (716) 263-1304
```

```
/ TELEFAX: (716) 263-1600
/ INFORMATION FOR SEQ ID NO: 19:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 107 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-895-914-19

Query Match      86.3%; Score 44; DB 3; Length 107;
Best Local Similarity 90.0%; Pred. No. 0.21;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 KASQDVGTSTV 10
Db      24 KASQDVGTAV 33

RESULT 9
US-09-357-710A-19
/ Sequence 19, Application US/09357710A
/ Patent No. 6290956
/ GENERAL INFORMATION:
/ APPLICANT: Bander, Neil H.
/ TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF PROSTATE CANCER
/ FILE REFERENCE: Lois M. Kwasiogoch: BZL 242/025
/ CURRENT APPLICATION NUMBER: US/09/357,710A
/ CURRENT FILING DATE: 1999-07-20
/ PRIOR APPLICATION NUMBER: US 08/838,682
/ PRIOR FILING DATE: 1997-04-09
/ PRIOR APPLICATION NUMBER: US 60/016,976
/ PRIOR FILING DATE: 1996-05-06
/ PRIOR APPLICATION NUMBER: US 60/022,125
/ PRIOR FILING DATE: 1996-07-18
/ NUMBER OF SEQ ID NOS: 21
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 19
/ LENGTH: 107
/ TYPE: PRT
/ ORGANISM: Mus sp.
US-09-357-710A-19

Query Match      86.3%; Score 44; DB 3; Length 107;
Best Local Similarity 90.0%; Pred. No. 0.21;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 KASQDVGTSTV 10
Db      24 KASQDVGTAV 33

RESULT 10
US-09-357-707-19
/ Sequence 19, Application US/09357707
/ Patent No. 6649163
/ GENERAL INFORMATION:
/ APPLICANT: Bander, Neil H.
/ TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CANCER
/ FILE REFERENCE: Lois M. Kwasiogoch: BZL 242/078
/ CURRENT APPLICATION NUMBER: US/09/357,707
/ CURRENT FILING DATE: 1999-07-20
/ PRIOR APPLICATION NUMBER: US 08/895,914
/ PRIOR FILING DATE: 1997-07-17
/ PRIOR APPLICATION NUMBER: US 08/838,682
/ PRIOR FILING DATE: 1997-04-09
/ PRIOR APPLICATION NUMBER: US 60/016,976
/ PRIOR FILING DATE: 1996-05-06
/ PRIOR APPLICATION NUMBER: US 60/022,125
/ PRIOR FILING DATE: 1996-07-18
/ NUMBER OF SEQ ID NOS: 21
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 19
```

```
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-357-707-19

Query Match      86.3%; Score 44; DB 4; Length 107;
Best Local Similarity 90.0%; Pred. No. 0.21;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 10
Db 24 KASQDVGTAV 33

RESULT 11
US-09-357-708-19
; Sequence 19, Application US/09357708
; Patent No. 6770450
; GENERAL INFORMATION:
; APPLICANT: Bandet, Neil H.
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CANCER
; FILE REFERENCE: Lois M. Kwaigiroch: BZL 242/028
; CURRENT APPLICATION NUMBER: US/09/357,708
; CURRENT FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: US 08/895,914
; PRIOR FILING DATE: 1997-07-17
; PRIOR APPLICATION NUMBER: US 08/838,682
; PRIOR FILING DATE: 1997-04-09
; PRIOR APPLICATION NUMBER: US 60/016,976
; PRIOR FILING DATE: 1996-05-06
; PRIOR APPLICATION NUMBER: US 60/022,125
; PRIOR FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-357-708-19

Query Match      86.3%; Score 44; DB 4; Length 107;
Best Local Similarity 90.0%; Pred. No. 0.21;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 10
Db 24 KASQDVGTAV 33

RESULT 12
US-08-797-689-18
; Sequence 18, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; COMPUTER: Macintosh

; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 249 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-18

Query Match      86.3%; Score 44; DB 2; Length 249;
Best Local Similarity 90.9%; Pred. No. 0.53;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11
Db 166 KASQDVDTSA 176

RESULT 13
US-09-984-186-18
; Sequence 18, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; COMPUTER: Macintosh
```

APPLICATION NUMBER: ER 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 249 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-984-186-18

Query Match 86.3%; Score 44; DB 4; Length 249;  
Best Local Similarity 90.9%; Pred. No. 0.53;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
Db 166 KASQDVDTSA 176

RESULT 14  
US-08-602-725-6  
Sequence 6, Application US/08602725  
Patent No. 5965710  
GENERAL INFORMATION:  
APPLICANT: BODMER, WALTER F  
APPLICANT: DURBIN, HELGA  
APPLICANT: SNARY, DAVID  
APPLICANT: STEWART, LORNA MD  
APPLICANT: YOUNG, SUSAN  
APPLICANT: BATES, PAUL A  
TITLE OF INVENTION: MONOCLONAL ANTIBODIES FOR USE IN  
DIAGNOSIS AND TREATMENT OF COLORECTAL CANCER  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,725  
FILING DATE: 02-FEB-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB94/01816  
FILING DATE: 19-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9317423  
FILING DATE: 21-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: SADOFF, B.J.  
REGISTRATION NUMBER: 36663  
REFERENCE/DOCKET NUMBER: 1090-8  
TELEPHONE: 703-816-4091  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: VL chain CDRI  
US-08-602-725-6

Query Match 84.3%; Score 43; DB 2; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.028;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
Db 1 KASQNVGTNA 11

RESULT 15  
US-08-646-265A-118  
Sequence 118, Application US/08646265A  
Patent No. 6214973  
GENERAL INFORMATION:  
APPLICANT: OHTOMO, Toshihiko  
APPLICANT: SATO, Koh  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
TITLE OF INVENTION: MEDULLOBLASTOMA CELLS  
NUMBER OF SEQUENCES: 132  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,265A  
FILING DATE: 09-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/JP94/01763  
FILING DATE: 19-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-291078  
FILING DATE: 19-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/184  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 118:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-265A-118

Query Match 84.3%; Score 43; DB 3; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.028;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTVA 11  
|||:||||  
Db 1 KASQVGTNVA 11

Search completed: August 9, 2005, 14:22:06  
Job time : 38 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:09:49 ; Search time 118.885 Seconds  
(without alignments)  
36.122 Million cell updates/sec

Title: US-10-089-452-27  
Perfect score: 51  
Sequence: 1 KASQDVGTSA 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1752860 seqs, 390397842 residues

Total number of hits satisfying chosen parameters: 1752860

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	11	9	US-09-253-794-20
2	51	100.0	11	16	US-10-680-734-20
3	51	100.0	11	17	US-10-755-382-20
4	51	100.0	106	9	US-09-253-794-4
5	51	100.0	106	9	US-09-253-794-19
6	51	100.0	106	16	US-10-680-734-4
7	51	100.0	106	16	US-10-680-734-19
8	51	100.0	106	17	US-10-755-382-4
9	51	100.0	106	17	US-10-755-382-19
10	51	100.0	115	14	US-10-270-073-4
11	51	100.0	232	13	US-10-006-771A-6

12	51	100.0	232	15	US-10-377-109-4	Sequence 4, Appli
13	51	100.0	256	14	US-10-270-073-6	Sequence 6, Appli
14	51	100.0	257	14	US-10-270-073-8	Sequence 8, Appli
15	51	100.0	261	14	US-10-270-071-18	Sequence 18, Appli
16	51	100.0	261	14	US-10-270-073-2	Sequence 2, Appli
17	51	100.0	261	14	US-10-328-190-6	Sequence 6, Appli
18	51	100.0	262	14	US-10-270-071-20	Sequence 20, Appli
19	51	100.0	262	14	US-10-270-071-24	Sequence 24, Appli
20	51	100.0	262	14	US-10-270-071-28	Sequence 28, Appli
21	51	100.0	262	14	US-10-270-071-34	Sequence 34, Appli
22	51	100.0	358	17	US-10-829-388-12	Sequence 12, Appli
23	51	100.0	363	17	US-10-829-388-2	Sequence 2, Appli
24	51	100.0	364	17	US-10-829-388-11	Sequence 11, Appli
25	51	100.0	370	17	US-10-829-388-1	Sequence 1, Appli
26	51	100.0	391	14	US-10-328-190-14	Sequence 14, Appli
27	51	100.0	443	13	US-10-006-773-2	Sequence 2, Appli
28	51	100.0	443	13	US-10-006-771A-2	Sequence 2, Appli
29	51	100.0	669	17	US-10-900-928-3	Sequence 3, Appli
30	48	94.1	11	14	US-10-281-479A-28	Sequence 28, Appli
31	48	94.1	11	14	US-10-275-180A-28	Sequence 28, Appli
32	48	94.1	11	14	US-10-286-132A-28	Sequence 28, Appli
33	48	94.1	11	16	US-10-297-371A-10	Sequence 10, Appli
34	48	94.1	108	10	US-09-215-163-21	Sequence 21, Appli
35	48	94.1	115	16	US-10-297-371A-8	Sequence 8, Appli
36	48	94.1	212	14	US-10-281-479A-72	Sequence 72, Appli
37	48	94.1	212	14	US-10-275-180A-72	Sequence 72, Appli
38	48	94.1	212	14	US-10-286-132A-72	Sequence 72, Appli
39	48	94.1	213	14	US-10-281-479A-46	Sequence 46, Appli
40	48	94.1	213	14	US-10-281-479A-73	Sequence 73, Appli
41	48	94.1	213	14	US-10-281-479A-74	Sequence 74, Appli
42	48	94.1	213	14	US-10-281-479A-75	Sequence 75, Appli
43	48	94.1	213	14	US-10-281-479A-76	Sequence 76, Appli
44	48	94.1	213	14	US-10-275-180A-46	Sequence 46, Appli
45	48	94.1	213	14	US-10-275-180A-73	Sequence 73, Appli

ALIGNMENTS

RESULT 1

US-09-253-794-20  
; Sequence 20, Application US/09253794  
; Patent No. US20020018750A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/253.794  
; FILING DATE: 22-Feb-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/318,157  
; FILING DATE: 05-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665  
; REFERENCE/DOCKET NUMBER: 18733/464  
; TELECOMMUNICATION INFORMATION:





/ REFERENCE/DOCKET NUMBER: 18733/464  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (202)672-5300  
/ TELEFAX: (202)672-5399  
/ TELEX: 904136  
/ INFORMATION FOR SEQ ID NO: 4:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 106 amino acids  
/ TYPE: amino acid  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
/ SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-253-794-4

Query Match 100.0%; Score 51; DB 9; Length 106;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
Db 24 KASQDVGTSA 34

## RESULT 5

US-09-253-794-19  
/ Sequence 19, Application US/09253794  
/ Patent No. US20020018750A1  
/ GENERAL INFORMATION:  
/ APPLICANT: ARMOUR, Kathryn L.  
/ TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
/ MOUSE MONOCLONAL ANTIBODIES

NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRES:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA

ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/253,794

FILING DATE: 22-Feb-1999  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/318,157  
FILING DATE: 05-OCT-1994

ATTORNEY/AGENT INFORMATION:  
NAME: SAXE, Bernhard D.  
REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 18733/464  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136

INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-253-794-19

Query Match 100.0%; Score 51; DB 9; Length 106;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
Db 24 KASQDVGTSA 34

## RESULT 6

US-10-680-734-4  
/ Sequence 4, Application US/10680734  
/ Publication No. US20040191248A1  
/ GENERAL INFORMATION:  
/ APPLICANT: GOLDENBERG, DAVID M.  
/ APPLICANT: HANSEN, HANS J.  
/ TITLE OF INVENTION: ANTIBODY THERAPY

/ FILE REFERENCE: 40923-0051US5  
/ CURRENT APPLICATION NUMBER: US/10/680,734  
/ CURRENT FILING DATE: 2003-10-08  
/ PRIOR APPLICATION NUMBER: 60/467,161  
/ PRIOR FILING DATE: 2003-05-02  
/ PRIOR APPLICATION NUMBER: PCT/US02/32307  
/ PRIOR FILING DATE: 2002-10-11  
/ PRIOR APPLICATION NUMBER: 60/416,531  
/ PRIOR FILING DATE: 2002-10-08  
/ NUMBER OF SEQ ID NOS: 27  
/ SOFTWARE: PatentIn Ver. 3.2  
/ SEQ ID NO 4

/ LENGTH: 106  
/ TYPE: PRT  
/ ORGANISM: Mus sp.  
US-10-680-734-4

Query Match 100.0%; Score 51; DB 16; Length 106;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
Db 24 KASQDVGTSA 34

## RESULT 7

US-10-680-734-19  
/ Sequence 19, Application US/10680734  
/ Publication No. US20040191248A1  
/ GENERAL INFORMATION:  
/ APPLICANT: GOLDENBERG, DAVID M.  
/ APPLICANT: HANSEN, HANS J.  
/ TITLE OF INVENTION: ANTIBODY THERAPY

/ FILE REFERENCE: 40923-0051US5  
/ CURRENT APPLICATION NUMBER: US/10/680,734  
/ CURRENT FILING DATE: 2003-10-08  
/ PRIOR APPLICATION NUMBER: 60/467,161  
/ PRIOR FILING DATE: 2003-05-02  
/ PRIOR APPLICATION NUMBER: PCT/US02/32307  
/ PRIOR FILING DATE: 2002-10-11  
/ PRIOR APPLICATION NUMBER: 60/416,531  
/ PRIOR FILING DATE: 2002-10-08  
/ NUMBER OF SEQ ID NOS: 27  
/ SOFTWARE: PatentIn Ver. 3.2  
/ SEQ ID NO 19

/ LENGTH: 106  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
US-10-680-734-19

Query Match 100.0%; Score 51; DB 16; Length 106;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
Db 24 KASQDVGTSA 34

## RESULT 8

US-10-755-382-4  
; Sequence 4, Application US/10755382  
; Publication No. US20050089538A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/755,382  
; FILING DATE: 13-Jan-2004  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/253,794  
; FILING DATE: 22-Feb-1999  
; APPLICATION NUMBER: US 08/318,157  
; FILING DATE: 05-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665  
; REFERENCE/DOCKET NUMBER: 18733/464  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 106 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-10-755-382-4

Query Match 100.0%; Score 51; DB 17; Length 106;  
Best Local Similarity 100.0%; Pred. No. 0.032; 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTGSA 11  
|||  
Db 24 KASQDVGTGSA 34

## RESULT 9

US-10-755-382-19  
; Sequence 19, Application US/10755382  
; Publication No. US20050089538A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA

ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/755,382  
; FILING DATE: 13-Jan-2004  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/253,794  
; FILING DATE: 22-Feb-1999  
; APPLICATION NUMBER: US 08/318,157  
; FILING DATE: 05-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665  
; REFERENCE/DOCKET NUMBER: 18733/464  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 106 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-10-755-382-19

Query Match 100.0%; Score 51; DB 17; Length 106;  
Best Local Similarity 100.0%; Pred. No. 0.032; 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTGSA 11  
|||  
Db 24 KASQDVGTGSA 34

## RESULT 10

US-10-270-073-4  
; Sequence 4, Application US/10270073  
; Publication No. US20030148409A1  
; GENERAL INFORMATION:  
; APPLICANT: ROSSI, EDMUND  
; APPLICANT: CHANG, CHIEN-HSING KEN  
; APPLICANT: GOLDENBERG, DAVID  
; TITLE OF INVENTION: DIRECT TARGETING BINDING PROTEINS  
; FILE REFERENCE: 042418/0112  
; CURRENT APPLICATION NUMBER: US/10/270,073  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: 60/328,835  
; PRIOR FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 60/341,881  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/345,641  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: 60/404,919  
; PRIOR FILING DATE: 2002-08-22  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: hmn-14VK deduced amino acid sequence  
US-10-270-073-4

Query Match 100.0%; Score 51; DB 14; Length 115;

```
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11
Db 24 KASQDVGTSA 34

RESULT 11
US-10-006-771A-6
; Sequence 6, Application US/10006771A
; Publication No. US20020165360A1
; GENERAL INFORMATION:
; APPLICANT: Junghans, Richard P.
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen
; FILE REFERENCE: 002
; CURRENT APPLICATION NUMBER: US/10/006,771A
; CURRENT FILING DATE: 2002-06-04
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens and Mus sp.
US-10-006-771A-6

Query Match 100.0%; Score 51; DB 13; Length 232;
Best Local Similarity 100.0%; Pred. No. 0.077;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11
Db 43 KASQDVGTSA 53

RESULT 12
US-10-377-109-4
; Sequence 4, Application US/10377109
; Publication No. US20040018557A1
; GENERAL INFORMATION:
; APPLICANT: QU, ZHENGXING
; APPLICANT: HANSEN, HANS
; APPLICANT: GOLDENBERG, DAVID
; TITLE OF INVENTION: BISPECIFIC ANTIBODY POINT MUTATIONS FOR ENHANCING RATE
; FILE REFERENCE: 018733/1159
; CURRENT APPLICATION NUMBER: US/10/377,109
; CURRENT FILING DATE: 2003-03-03
; PRIOR FILING DATE: 2003-03-03
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-377-109-4

Query Match 100.0%; Score 51; DB 15; Length 232;
Best Local Similarity 100.0%; Pred. No. 0.077;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11
Db 43 KASQDVGTSA 53

RESULT 13
US-10-270-073-6
; Sequence 6, Application US/10270073
; Publication No. US20030148409A1
```

```
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND
; APPLICANT: CHANG, CHIEN-HSING KEN
; APPLICANT: GOLDENBERG, DAVID
; TITLE OF INVENTION: DIRECT TARGETING BINDING PROTEINS
; FILE REFERENCE: 042418/0112
; CURRENT APPLICATION NUMBER: US/10/270,073
; CURRENT FILING DATE: 2002-10-15
; PRIOR FILING DATE: 2001-10-15
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/341,881
; PRIOR APPLICATION NUMBER: 60/345,641
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: 60/404,919
; PRIOR FILING DATE: 2002-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: hmn-14-0 amino acid sequence
US-10-270-073-6

Query Match 100.0%; Score 51; DB 14; Length 256;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11
Db 165 KASQDVGTSA 175

RESULT 14
US-10-270-073-8
; Sequence 8, Application US/10270073
; Publication No. US20030148409A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND
; APPLICANT: CHANG, CHIEN-HSING KEN
; APPLICANT: GOLDENBERG, DAVID
; TITLE OF INVENTION: DIRECT TARGETING BINDING PROTEINS
; FILE REFERENCE: 042418/0112
; CURRENT APPLICATION NUMBER: US/10/270,073
; CURRENT FILING DATE: 2002-10-15
; PRIOR FILING DATE: 2001-10-15
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/341,881
; PRIOR APPLICATION NUMBER: 60/345,641
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: 60/404,919
; PRIOR FILING DATE: 2002-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: hmn-14-IG amino acid sequence
US-10-270-073-8

Query Match 100.0%; Score 51; DB 14; Length 257;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11
Db 165 KASQDVGTSA 175
```

Db 166 KASQDVGTSA 176

## RESULT 15

US-10-270-071-18  
; Sequence 18, Application US/10270071  
; Publication No. US2003011333A1  
; GENERAL INFORMATION:  
; APPLICANT: ROSSI, EDMUND  
; APPLICANT: CHANG, CHIEN-HSING KEN  
; TITLE OF INVENTION: AFFINITY ENHANCEMENT AGENTS  
; FILE REFERENCE: 042418/0110  
; CURRENT APPLICATION NUMBER: US/10/270,071  
; CURRENT FILING DATE: 2003-01-21  
; PRIOR APPLICATION NUMBER: 60/328,835  
; PRIOR FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 60/341,881  
; PRIOR FILING DATE: 2001-12-21  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide  
; OTHER INFORMATION: sequence of HMN14VK  
US-10-270-071-18

Query Match 100.0%; Score 51; DB 14; Length 261;  
Best Local Similarity 100.0%; Pred. No. 0.088;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
| | | | | | | | | |  
Db 170 KASQDVGTSA 180

Search completed: August 9, 2005, 14:26:49  
Job time : 119.885 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:06:08 ; Search time 24.5385 Seconds  
(without alignments)  
43.132 Million cell updates/sec

Title: US-10-089-452-27

Perfect score: 51

Sequence: 1 KASQDVGTSA 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79:\*

1: piri:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	94.1	117	2	S42466
2	43	84.3	82	2	S38560
3	43	84.3	90	2	S38561
4	42	82.4	93	2	S38559
5	42	82.4	107	2	S09967
6	42	82.4	107	2	S32191
7	42	82.4	107	2	S32192
8	42	82.4	119	2	PQ0265
9	42	82.4	149	1	KVMS11
10	39	76.5	88	2	A37262
11	38	74.5	86	2	C28195
12	38	74.5	106	2	PL0088
13	38	74.5	107	2	B28195
14	38	74.5	107	2	A28195
15	38	74.5	128	2	A47159
16	37	72.5	98	2	PH1073
17	37	72.5	110	2	D72673
18	37	72.5	115	2	A25924
19	37	72.5	127	2	S04577
20	37	72.5	152	2	S30751
21	37	72.5	254	2	E69853
22	37	72.5	520	2	T19661
23	36	70.6	87	2	PH1082
24	36	70.6	88	2	PL0260
25	36	70.6	105	2	PL0260
26	36	70.6	115	1	KVMSL7
27	36	70.6	124	2	T40071
28	36	70.6	128	2	PN0445
29	36	70.6	130	1	KVMSM4

30 36 70.6 317 2 T33277  
31 36 70.6 572 1 HNN2B3  
32 36 70.6 737 2 A87626  
33 35 68.6 94 2 F33730  
34 35 68.6 165 2 JC5737  
35 35 68.6 250 2 G81447  
36 35 68.6 298 2 B75429  
37 34 66.7 91 2 PH1071  
38 34 66.7 96 2 PH1070  
39 34 66.7 98 2 PH1072  
40 34 66.7 100 2 H38601  
41 34 66.7 106 2 PL0259  
42 34 66.7 107 2 PL0268  
43 34 66.7 108 2 PL0083  
44 34 66.7 108 2 C30502  
45 34 66.7 108 2 PL0204

#### ALIGNMENTS

##### RESULT 1

S42466

Ig kappa chain V region - mouse

C;Species: Mus musculus (house mouse)

C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 21-Jan-2000

C;Accession: S42466

R;Shivanov, P.A.; Beapalov, I.A.; Terletekaya, H.N.; Deyev, S.M.

submitted to the EMBL Data Library, March 1994

A;Reference number: S42466

A;Accession: S42466

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-117 <SH1>

A;Cross-references: EMBL:X78108; NID:g460824; PIDN:CAA54998.1; PID:g460825

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotrimer; immunoglobulin

F;26-100/Domain: immunoglobulin homology <IMV>

Query Match 94.1%; Score 48; DB 2; Length 117;  
Best Local Similarity 90.9%; Pred. No. 0.019;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11

Db 34 KASQDVGTAVA 44

##### RESULT 2

S38560

Ig light chain V region (ASWBL) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999

C;Accession: S38560

R;Monestier, M.; Losman, L.J.; Novick, K.E.; Aris, J.P.

submitted to the EMBL Data Library, September 1993

A;Description: Molecular analysis of mercury-induced anti-nucleolar antibodies in H-2s M

A;Reference number: S38559

A;Accession: S38560

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-82 <MON>

A;Cross-references: EMBL:X75102; NID:g414147; PIDN:CAA52993.1; PID:g414148

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: immunoglobulin

Query Match 84.3%; Score 43; DB 2; Length 82;  
Best Local Similarity 81.8%; Pred. No. 0.14;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11

Db 24 KASQNVGTAVA 34

```
RESULT 3
S38561
IG kappa chain V region (ASWPL) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S38561
R:Monestier, M.; Lozman, L.J.; Novick, K.E.; Aris, J.P.
submitted to the EMBL Data Library, September 1993
A:Description: Molecular analysis of mercury-induced anti-nucleolar antibodies in H-2s M
A:Reference number: S38559
A:Accession: S38561
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-90 <MON>
A:Cross-references: EMBL:X75103; NID:g414155; PIDN:CAAS2994.1; PID:g414156
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin
F:16-90/Domain: immunoglobulin homology <IMM>

Query Match      84.3%; Score 43; DB 2; Length 90;
Best Local Similarity 81.8%; Pred. No. 0.15;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSTVA 11
   |||||:|:|
Db 24 KASQNVGTAVA 34

RESULT 4
S38559
IG light chain V region (ASWAL) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C:Accession: S38559
R:Monestier, M.; Lozman, L.J.; Novick, K.E.; Aris, J.P.
submitted to the EMBL Data Library, September 1993
A:Description: Molecular analysis of mercury-induced anti-nucleolar antibodies in H-2s M
A:Reference number: S38559
A:Accession: S38559
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-93 <MON>
A:Cross-references: EMBL:X75101; NID:g414145; PIDN:CAAS2992.1; PID:g414146
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin
F:16-90/Domain: immunoglobulin homology <IMM>

Query Match      82.4%; Score 42; DB 2; Length 93;
Best Local Similarity 81.8%; Pred. No. 0.25;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KASQDVGTSTVA 11
   |||||:|:|
Db 24 KASQDVSTAVA 34

RESULT 5
S09967
IG kappa chain V-J region (31-9D) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 16-Aug-1996
C:Accession: S09967
R:Reininger, L.; Shibata, T.; Ozaki, S.; Shirai, T.; Jatton, J.C.; Izui, S.
Eur. J. Immunol. 20, 771-777, 1990
A:Title: Variable region sequences of pathogenic anti-mouse red blood cell autoantibodies
A:Reference number: S09955; MUID:90269328; PMID:2347362
A:Accession: S09967
A:Molecule type: mRNA
A:Residues: 1-107 <REI>
A:Cross-references: EMBL:X51855
C:Superfamily: immunoglobulin V region; immunoglobulin homology
```

C:Keywords: heterotetramer; immunoglobulin

Query Match 82.4%; Score 42; DB 2; Length 107;  
Best Local Similarity 81.8%; Pred. No. 0.29;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KASQDVGTSTVA 11  
 |||||:|:|  
Db 24 KASQDVSTAVA 34

```
RESULT 6
S32191
IG kappa chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 21-Jan-2000
C:Accession: S32191
R:Izui, S.
submitted to the EMBL Data Library, February 1993
A:Reference number: S32185
A:Accession: S32191
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-107 <IZU>
A:Cross-references: EMBL:X70095; NID:g288260; PIDN:CAA49700.1; PID:g288261
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-90/Domain: immunoglobulin homology <IMM>
```

Query Match 82.4%; Score 42; DB 2; Length 107;  
Best Local Similarity 81.8%; Pred. No. 0.29;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KASQDVGTSTVA 11  
 |||||:|:|  
Db 24 KASQDVSTAVA 34

```
RESULT 7
S32192
IG kappa chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 21-Jan-2000
C:Accession: S32192
R:Izui, S.
submitted to the EMBL Data Library, February 1993
A:Reference number: S32185
A:Accession: S32192
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-107 <IZU>
A:Cross-references: EMBL:X70097; NID:g288262; PIDN:CAA49701.1; PID:g288263
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-90/Domain: immunoglobulin homology <IMM>
```

Query Match 82.4%; Score 42; DB 2; Length 107;  
Best Local Similarity 81.8%; Pred. No. 0.29;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KASQDVGTSTVA 11  
 |||||:|:|  
Db 24 KASQDVSTAVA 34

```
RESULT 8
PQ0265
IG kappa chain V region (MC1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Jan-2000
C:Accession: PQ0265
R:Lohman, K.L.; Carrillo, M.A.; Kennedy, R.C.
Gene 105, 283-284, 1991
```

A;Title: Sequence analysis of the variable region of a mouse gene encoding a monoclonal  
A;Reference number: PQ0265; MUID:92039046; PMID:1937027  
A;Accession: PQ0265  
A;Molecule type: mRNA  
A;Residues: 1-119 <LOH>  
A;Cross-references: GB:M59985  
A;Comment: This protein recognizes a restricted idiotype associated with antibodies spec  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;28-102/Domain: immunoglobulin homology <IMM>  
F;36-46/Region: complementarity-determining 1  
F;62-68/Region: complementarity-determining 2  
F;101-109/Region: complementarity-determining 3

Query Match 82.4%; Score 42; DB 2; Length 119;  
Best Local Similarity 81.8%; Pred. No. 0.33;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11  
||||| :||  
Db 36 KASQDVSTTVA 46  
||||| :||

RESULT 9  
KVMS11  
Ig kappa chain precursor V region (MPC11) - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 19-Feb-1984 #sequence\_revision 19-Feb-1984 #text\_change 09-Jul-2004  
C;Accession: A90823; A90753; A90298; A01916  
R;Kelley, D.E.; Coleclough, C.; Perry, R.P.  
Cell 29, 681-689, 1982  
A;Title: Functional significance and evolutionary development of the 5'-terminal regions  
A;Reference number: A90823; MUID:83001944; PMID:6288267  
A;Accession: A90823  
A;Molecule type: DNA  
A;Residues: 1-71 <REL>  
A;Cross-references: UNIPROT:P01633  
A;Note: the sequence was determined from the differentiated gene  
R;Rabbits, T.H.; Hanlyn, P.H.; Matthysens, G.; Roe, B.A.  
Can. J. Biochem. 58, 176-187, 1980  
A;Title: The variability, arrangement, and rearrangement of immunoglobulin genes.  
A;Reference number: A90753; MUID:80176554; PMID:6245773  
A;Accession: A90753  
A;Molecule type: mRNA  
A;Residues: 41-149 <RAB>  
R;Smith, G.P.  
Biochem. J. 171, 337-347, 1978  
A;Title: Sequence of the full-length immunoglobulin kappa-chain of mouse myeloma MPC 11.  
A;Reference number: A90298; MUID:78186617; PMID:418775  
A;Contents: myeloma protein MPC11  
A;Accession: A90298  
A;Molecule type: protein  
A;Residues: 30-149 <SMI>  
A;Note: the amidation states of residues 58, 69, 101, 111, 120, 122, and 123 were not de  
C;Comment: The mature chain has 12 additional residues at its amino end, due to a tandem  
42 corresponds to the amino-terminal residue of typical kappa chains.  
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: duplication; heterotetramer; immunoglobulin  
F;1-29/Domain: signal sequence #status predicted <SIG>  
F;30-149/Product: Ig kappa chain V region (MPC11) #status experimental <MAT>  
F;57-133/Domain: immunoglobulin homology <IMM>  
F;64-129/Disulfide bonds: #status predicted

Query Match 82.4%; Score 42; DB 1; Length 149;  
Best Local Similarity 81.8%; Pred. No. 0.41;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11  
||||| :||  
Db 65 KASQDVSTTVA 75  
||||| :||

## RESULT 10

A37262  
Ig kappa chain V region (6H2) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Aug-1991 #sequence\_revision 03-Apr-1992 #text\_change 09-Jul-2004  
C;Accession: A37262  
R;Goshorn, S.C.; Retzel, E.; Jermerson, R.  
J. Biol. Chem. 266, 2134-2142, 1991  
A;Title: Common structural features among monoclonal antibodies binding the same antigen  
A;Reference number: A38601; MUID:91115823; PMID:1703527  
A;Accession: A37262  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-88 <GOS>  
A;Cross-references: UNIPROT:Q8VIJ0; GB:M57987  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin

Query Match 76.5%; Score 39; DB 2; Length 88;  
Best Local Similarity 72.7%; Pred. No. 0.99;  
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11  
||||| :||  
Db 4 KASQNVGSNVA 14  
||||| :||

## RESULT 11

C28195  
Ig kappa chain V region (anti-haloperidol antibody C) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 09-May-1997  
C;Accession: C28195  
R;Sherman M.A.; Deans, R.J.; Bolger, M.B.  
J. Biol. Chem. 263, 4059-4063, 1988  
A;Title: Haloperidol binding to monoclonal antibodies. Hypervariable region amino acid s  
A;Reference number: A28195; MUID:88153717; PMID:3267217  
A;Accession: C28195  
A;Molecule type: mRNA  
A;Residues: 1-86 <SHE>  
A;Cross-references: GB:M19768  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin

Query Match 74.5%; Score 38; DB 2; Length 86;  
Best Local Similarity 72.7%; Pred. No. 1.5;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KASQDVGTSA 11  
||||| :||  
Db 3 KASQNVGNVA 13  
||||| :||

## RESULT 12

PL0088  
Ig kappa chain V regions (12S18-1, 12S28-16) - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 21-Jan-2000  
C;Accession: PL0088  
R;Meek, K.; Haemann, C.; Pollok, B.; Alkan, S.S.; Brait, M.; Slaoui, M.; Urbain, J.; Ca  
J. Exp. Med. 169, 519-533, 1989  
A;Title: Structural characterization of antiidiotypic antibodies; evidence that Ab2s are  
A;Reference number: PL0080; MUID:89094248; PMID:2492056  
A;Accession: PL0088  
A;Molecule type: mRNA  
A;Residues: 1-106 <MBE>  
A;Cross-references: GB:X58581; GB:Y00794; NID:951592; PIDN:CAA41457.1; PID:9930151  
A;Note: the sequence shown here is from the V kappa regions an antiidiotypic monoclonal  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;16-89/Domain: immunoglobulin homology <IMM>

Query Match 74.5%; Score 38; DB 2; Length 106;  
Best Local Similarity 72.7%; Pred. No. 1.9;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
DB 24 KASQVGNVA 34  
|||||:|

## RESULT 13

B28195  
Ig kappa chain V region (anti-haloperidol antibody B) - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 21-Jan-2000  
C:Accession: B28195  
R:Sherman, M.A.; Deans, R.J.; Bolger, M.B.  
J. Biol. Chem. 263, 4059-4063, 1988  
A:Title: Haloperidol binding to monoclonal antibodies. Hypervariable region amino acid  
A:Reference number: A28195; MUID:88153717; PMID:3267217  
A:Accession: B28195  
A:Molecule type: mRNA  
A:Residues: 1-107 <SHE>  
A:Cross-references: EMBL:M19767; NID:g197041; PIDN:AAA38892.1; PID:g197042  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:16-90/Domain: immunoglobulin homology <IMM>

Query Match 74.5%; Score 38; DB 2; Length 107;  
Best Local Similarity 72.7%; Pred. No. 1.9;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
DB 24 KASQVGNVA 34  
|||||:|

## RESULT 14

A28195  
Ig kappa chain V region (anti-haloperidol antibody A) - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 21-Jan-2000  
C:Accession: A28195  
R:Sherman, M.A.; Deans, R.J.; Bolger, M.B.  
J. Biol. Chem. 263, 4059-4063, 1988  
A:Title: Haloperidol binding to monoclonal antibodies. Hypervariable region amino acid  
A:Reference number: A28195; MUID:88153717; PMID:3267217  
A:Accession: A28195  
A:Molecule type: mRNA  
A:Residues: 1-107 <SHE>  
A:Cross-references: GB:M19766; NID:g197039; PIDN:AAA38891.1; PID:g197040  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:16-90/Domain: immunoglobulin homology <IMM>

Query Match 74.5%; Score 38; DB 2; Length 107;  
Best Local Similarity 72.7%; Pred. No. 1.9;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
DB 24 KASQVGNVA 34  
|||||:|

## RESULT 15

A47159  
Ig lambda chain V region (CEA-specific maId T84.66) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000  
C:Accession: A47159  
R:Gaide, F.J.; Pieper, D.; Roder, U.W.; Shively, J.E.; Wagener, C.; Neumaier, M.  
J. Biol. Chem. 268, 14138-14145, 1993  
A:Title: Molecular characterization of a cloned idiotypic cascade containing a network  
A:Reference number: A47159; MUID:93300804; PMID:7686150

A:Accession: A47159  
A>Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-128 <GAI>  
A:Experimental source: hybridoma 6G6.C4  
A>Note: sequence extracted from NCBI backbone (NCBI:134419, NCBIP:134420)  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:36-110/Domain: immunoglobulin homology <IMM>

Query Match 74.5%; Score 38; DB 2; Length 128;  
Best Local Similarity 72.7%; Pred. No. 2.3;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KASQDVGTSA 11  
DB 44 KASQVGNVA 54  
|||||:|

Search completed: August 9, 2005, 14:20:38  
Job time : 25.5385 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:05:28 ; Search time 118.038 Seconds  
(without alignments)  
47.721 Million cell updates/sec

Title: US-10-089-452-27  
Perfect score: 51  
Sequence: 1 KASQDVGTSVA 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	84.3	108	2 Q8VIJ0	Q8VIJ0 mus musculus
2	42	82.4	149	1 KV5A MOUSE	P01633 mus musculus
3	37	72.5	110	2 Q9YD79	Q9YD79 aeropyrum p
4	37	72.5	254	2 Q34578	Q34578 bacillus su
5	37	72.5	409	2 Q8JZV7	Q8JZV7 mus musculus
6	37	72.5	409	2 Q8BK10	Q8BK10 mus musculus
7	37	72.5	520	2 Q18361	Q18361 caenorhabdi
8	36	70.6	115	1 KV51 MOUSE	P01642 mus musculus
9	36	70.6	124	1 CV51 SCHPO	Q94391 schizosacch
10	36	70.6	130	1 KV5G MOUSE	P01639 mus musculus
11	36	70.6	317	2 Q61933	Q61933 caenorhabdi
12	36	70.6	572	1 HEMA P13B	P06167 bovine para
13	36	70.6	572	2 Q65689	Q65689 bovine para
14	36	70.6	737	1 CATA CAUCR	Q1066 caulobacter
15	36	70.6	2598	2 Q7RNG6	Q7RNG6 plasmodium
16	35	68.6	83	2 Q6Z997	Q6Z997 oryza sativ
17	35	68.6	165	1 ATPF STRMU	P95785 streptococc
18	35	68.6	250	2 Q9PIK7	Q9PIK7 campylobact
19	35	68.6	298	1 UB72 DEIRA	Q9RV58 deinococcus
20	35	68.6	311	2 Q9TNJ1	Q9TNJ1 haplomitriu
21	35	68.6	374	2 Q8CP46	Q8CP46 staphylococ
22	35	68.6	661	2 Q8R0V4	Q8R0V4 mus musculus
23	35	68.6	801	2 Q8CN69	Q8CN69 staphylococ
24	35	68.6	930	2 Q9DBM1	Q9DBM1 mus musculus
25	35	68.6	1158	2 Q74ZC6	Q74ZC6 ashbya goss
26	34	66.7	83	2 Q8CEV4	Q8CEV4 mus musculus
27	34	66.7	158	1 GRE4 PSESM	Q87WP5 pseudomonas
28	34	66.7	182	2 Q8TU86	Q8TU86 methanosaar
29	34	66.7	197	2 Q7QCN2	Q7QCN2 anopheles g
30	34	66.7	212	2 Q82CD8	Q82CD8 streptomyce
31	34	66.7	218	2 Q7WLF0	Q7WLF0 bordetella

32	34	66.7	220	2 Q7VX51	Q7VX51 bordetella
33	34	66.7	220	2 Q7W806	Q7W806 bordetella
34	34	66.7	251	1 RS3B SCHPO	Q94438 schizosacch
35	34	66.7	295	2 Q98D96	Q98D96 rhizobium l
36	34	66.7	319	2 Q741L9	Q741L9 mycobacteri
37	34	66.7	326	2 Q9LTW7	Q9LTW7 arabidopsis
38	34	66.7	328	2 Q9C7B5	Q9C7B5 arabidopsis
39	34	66.7	330	2 Q6RW08	Q6RW08 uncultured
40	34	66.7	336	2 Q8NCQ7	Q8NCQ7 homo sapien
41	34	66.7	395	2 Q8EQN5	Q8EQN5 oceanobacil
42	34	66.7	442	2 Q6MWC2	Q6MWC2 oryza sativ
43	34	66.7	463	2 Q7EZC9	Q7EZC9 oryza sativ
44	34	66.7	524	2 Q941Q3	Q941Q3 arabidopsis
45	34	66.7	565	2 Q51505	Q51505 pseudomonas

#### ALIGNMENTS

##### RESULT 1

Q8VIJ0 Q8VIJ0 PRELIMINARY; PRT; 108 AA.  
AC Q8VIJ0;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Anti-DNA light chain (Fragment).  
GN Name=Vk19;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C3H/HeJ-lpr/lpr;  
RX MEDLINE=86409289; PubMed=8814271;  
RA Wloch M.K., Alexander A.L., Phippen A.M., Pisetsky D.S., Gilkeson G.S.;  
RT "Differences in V kappa gene utilization and VH CDR3 sequence among  
anti-DNA from C3H-lpr mice and lupus mice with nephritis.";  
RL Eur. J. Immunol. 26:2225-2233(1996).  
DR EMBL; U59155; BAB02917.1; -.  
DR PIR; A33933; A33933.  
DR PIR; A37262; A37262.  
DR PIR; PH1072; PH1072.  
DR HSSP; P01634; IIGC.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
FT NON\_TER 1  
FT NON\_TER 108  
SQ SEQUENCE 108 AA; 11859 MW; 68506D75613DBFBE CRC64;

Query Match 84.3%; Score 43; DB 2; Length 108;

Best Local Similarity 81.8%; Pred. No. 0.75; Mismatches 2; Indels 0; Gaps 0;  
Matches 9; Conservative 2;

Qv 1 KASQDVGTSVA 11

Db 24 KASQNVGTNVA 34

##### RESULT 2

KV5A MOUSE  
ID KV5A MOUSE STANDARD; PRT; 149 AA.  
AC P01633;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Ig kappa chain V-V region MPC11 precursor.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]

```
RP SEQUENCE OF 1-71 FROM N.A.
RX MEDLINE=83001944; PubMed=6288267; DOI=10.1016/0092-8674(82)90184-2;
RA Kelley D.E., Coleclough C., Perry R.P.;
RT "Functional significance and evolutionary development of the 5'-
RL terminal regions of immunoglobulin variable-region genes.";
RN Cell 29:681-689(1982).
RN [2]
RP SEQUENCE OF 41-149 FROM N.A.
RX MEDLINE=80176554; PubMed=6245773;
RA Rabbitts T.H., Hanly P.H., Matthysens G., Roe B.A.;
RT "The variability, arrangement, and rearrangement of immunoglobulin
RL genes.";
RN Can. J. Biochem. 58:176-187(1980).
RN [3]
RP SEQUENCE OF 30-149.
RX MEDLINE=78186617; PubMed=418775;
RA Smith G.P.;
RT "Sequence of the full-length immunoglobulin kappa-chain of mouse
RL myeloma MPC 11.";
RN Biochem. J. 171:337-347(1978).
CC -!- MISCELLANEOUS: The mature chain has 12 additional residues at its
CC amino end, due to a tandem duplication of 36 nucleotides after the
CC codon for residue 36. Residue 42 corresponds to the amino-terminal
CC residue of typical kappa chains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00561; AAA38776.1; -.
DR PIR; A90823; KWS11.
DR HSP; P01634; IIGC.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR Direct protein sequencing; Immunoglobulin V region; Repeat; Signal.
KW SIGNAL 1 29
FT CHAIN 30 149 Ig kappa chain V-V region MPC11.
FT DOMAIN 42 64 Framework-1.
FT DOMAIN 65 75 Complementarity-determining-1.
FT DOMAIN 76 90 Framework-2.
FT DOMAIN 91 97 Complementarity-determining-2.
FT DOMAIN 98 129 Framework-3.
FT DOMAIN 130 138 Complementarity-determining-3.
FT DOMAIN 139 148 Framework-4.
FT REPEAT 26 35
FT REPEAT 38 47
FT NON_TER 149 149
SQ SEQUENCE 149 AA; 16434 MW; B0480C87B682AC3E CRC64;
```

```
Query Match 82.4%; Score 42; DB 1; Length 149;
Best Local Similarity 81.8%; Pred. No. 1.7;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 KASQDVGTGSA 11

Db 65 KASQDVGTGSA 75

RESULT 3

Q9YDV9

ID Q9YDV9 PRELIMINARY; PRT; 110 AA.

AC Q9YDV9;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Hypothetical protein APE0810.

```
GN OrderedLocusNames=APE0810;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1.
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA79788.1; -.
DR PIR; D72673; D72673.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 110 AA; 12572 MW; D1A1898896CF8F34 CRC64;

Query Match 72.5%; Score 37; DB 2; Length 110;
Best Local Similarity 72.7%; Pred. No. 14;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KASQDVGTGSA 11
Db 65 KLSQDLGQSV 75

RESULT 4
O34578 PRELIMINARY; PRT; 254 AA.
AC O34578; Q796M8;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein YJNA.
GN Name=YJNA; OrderedLocusNames=BSU12400;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98240225; PubMed=9579062;
RA Rivolta C., Soldo B., Lazarevic V., Joris B., Mael C., Karamata D.;
RT "A 35.7 kb DNA fragment from the Bacillus subtilis chromosome
RT containing a putative 12.3 kb operon involved in hexuronate catabolism
RT and a perfectly symmetrical hypothetical catabolite-responsive
RT element.";
RL Microbiology 144:877-884(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Bortiss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mael C.,
```

RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,  
 RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,  
 RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,  
 RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,  
 RA Rey M., Reynolds S., Rieger M., Rivola C., Rocha E., Roche B.,  
 RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,  
 RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,  
 RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,  
 RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,  
 RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,  
 RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,  
 RA Weitzneger T., Winters P., Wipat A., Yamamoto H., Yamane K.,  
 RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,  
 RA Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.";  
 RL Nature 390:249-256(1997).  
 DR EMBL; AF015825; AAC46336.1; -;  
 DR EMBL; Z99110; CAB13097.1; -;  
 DR PIR; E69853; E69853.  
 DR InterPro; IPR002781; DUF81.  
 DR Pfam; PF01925; DUF81; 1.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 254 AA; 27135 MW; C1513F06D214EC3D CRC64;

Query Match 72.5%; Score 37; DB 2; Length 254;  
 Best Local Similarity 72.7%; Pred. No. 34;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KASQDVGTGVA 11  
 |||||  
 DB 178 KASQIVGTGDI 188

RESULT 5  
 Q8JZV7 PRELIMINARY; PRT; 409 AA.  
 AC 08JZV7  
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)  
 DE RIKEN cDNA 5730457F11 (Mus musculus cDNA, RIKEN full-length enriched  
 DE library, clone:G430059K10 product:weakly similar to PUTATIVE N-  
 DE ACETYLGLUCOSAMINE-6-PHOSPHATE DEACETYLASE) (EC 3.5.1.25) (GLCNAC 6-P  
 DE DEACETYLASE).  
 CN Name=5730457F11Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Liver;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalske U., Smalley D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]  
 RL SEQUENCE FROM N.A.  
 RN STRAIN=FVB/N; TISSUE=Liver;  
 RC Strausberg R.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21085650; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 [5]  
 RP SEQUENCE FROM N.A.  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 [8]  
 RP SEQUENCE FROM N.A.  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohazato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC037005; AAH37005.1; -;  
 DR EMBL; AK090009; BAC41042.1; -;  
 DR HSSP; O34450; 1UN7.  
 DR GSD; MGI:2443978; 5730457F11Rik.  
 DR GO; GO:0016787; F-hydrolase activity; IEA.  
 DR GO; GO:0008448; F:N-acetylglucosamine-6-phosphate deacetylase. . .; IEA.  
 DR GO; GO:0006044; P:N-acetylglucosamine metabolism; IEA.  
 DR InterPro; IPR006680; P-hydrolase.  
 DR InterPro; IPR011550; Amidohydro\_1.  
 DR InterPro; IPR011059; Metallo\_hydrolase.  
 DR InterPro; IPR003764; Naga.  
 DR Pfam; PF01979; Amidohydro\_1; 1.  
 DR ProDom; PD001248; Amidohydro\_like; 1.





```
DR PRINTS; PR00363; CYTOCHROME_B5.
DR ProDom; PD000812; CYC_B5; 1.
DR PROSITE; PS0255; CYTOCHROME_B5_2; 1.
KW Electron transport; Heme; Iron; Microsome; Transmembrane.
FT TRANSMEM 100 120 Potential.
FT METAL 38 38 Iron (heme axial ligand) (By similarity).
FT METAL 62 62 Iron (heme axial ligand) (By similarity).
SQ SEQUENCE 124 AA; 13910 MW; 299D12CB693831D4 CRC64;

Query Match
Best Local Similarity 70.6%; Score 36; DB 1; Length 124;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KASQDVGTSTVA 11
|||:|:|:|
DB 55 KAYQDIGHSIA 65

RESULT 10
KV5G_MOUSE STANDARD; PRT; 130 AA.
AC P01639; P01640;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ig kappa chain V-V region MOPC 41 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RX MEDLINE=79221900; PubMed=111146;
RA Seidman J.G., Max E.B., Leder P.;
RT "A kappa-immunoglobulin gene is formed by site-specific recombination
RT without further somatic mutation.";
RL Nature 280:370-375(1979).
RN [2]
RP SEQUENCE OF 1-33.
RX MEDLINE=77148916; PubMed=403522;
RA Burstein Y., Schechter I.;
RT "Amino acid sequence of the NH2-terminal extra piece segments of the
RT precursors of mouse immunoglobulin lambda1-type and kappa-type light
RT chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 74:716-720(1977).
RN [3]
RP SEQUENCE OF 23-130.
RX MEDLINE=67056897; PubMed=4162931;
RA Gray W.R., Dreyer W.J., Hood L.;
RT "Mechanism of antibody synthesis: size differences between mouse kappa
RT chains.";
RL Science 155:465-467(1967).
CC -!- MISCELLANEOUS: This precursor was synthesized in a cell-free
CC system directed by mRNA isolated from myeloma polysomes.
CC -!- MISCELLANEOUS: This is a Bence-Jones protein.
DR PIR; A93211; KMSM4.
DR HSP; P01594; IJVS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
KW Bence-Jones protein; Direct protein sequencing;
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 22
FT CHAIN 23 130 Ig kappa chain V-V region MOPC 41.
FT DOMAIN 23 45 Framework-1.
FT DOMAIN 23 45 Complementarity-determining-1.
FT DOMAIN 46 56 Framework-2.
FT DOMAIN 57 71 Framework-3.
FT DOMAIN 72 78 Complementarity-determining-2.
FT DOMAIN 79 110 Framework-4.
FT DOMAIN 111 119 Complementarity-determining-3.
FT DOMAIN 120 129 Framework-4.

FT DISULFID 45 110 By similarity.
FT VARIANT 1 2 Missing (in 25% of the molecules).
FT NON_TER 130 130
SQ SEQUENCE 130 AA; 14311 MW; 5EPE0FE71D5F1BEC CRC64;

Query Match
Best Local Similarity 70.6%; Score 36; DB 1; Length 130;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KASQDVGTSTV 10
:|:|:|:|:|
DB 46 RASQDIGSSL 55

RESULT 11
O61933 PRELIMINARY; PRT; 317 AA.
AC O61933;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein R09A1.3.
GN Name=R09A1.3; ORFNames=R09A1.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Bentley D., Kemp K., Harper M.;
RT "The sequence of C. elegans cosmid R09A1.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF068711; AAC17777.1; -.
DR PIR; T33277; T33277.
DR WormBase; WBGene00019973; R09A1.3.
DR WormPep; R09A1.3; CE18129.
DR Hypothetical protein.
SQ SEQUENCE 317 AA; 35559 MW; 87127052557A517D CRC64;

Query Match
Best Local Similarity 70.6%; Score 36; DB 2; Length 317;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KASQDVGTSTVA 11
|:|:|:|:|
DB 37 KESMDIGTSTA 47

RESULT 12
HENA_P13B STANDARD; PRT; 572 AA.
ID HENA_P13B
AC P06167;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
```

```
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hemagglutinin-neuraminidase (EC 3.2.1.18) .
GN Name=HN;
OS Bovine parainfluenza 3 virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Respirovirus.
OX NCBI_TaxID=11215;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=910N;
RX MEDLINE=87174819; PubMed=3031615;
RA Suzu S., Sakai Y., Shioda T., Shibuta H.;
RT "Nucleotide sequence of the bovine parainfluenza 3 virus genome: the
RT genes of the F and HN glycoproteins."
RL Nucleic Acids Res. 15:2945-2958(1987) .
CC -I- FUNCTION: Hemagglutinin is responsible for attaching the virus to
CC cell receptors and for initiating infection. Neuraminidase
CC activity helps the efficient spread of the virus by dissociating
CC the mature virions from the neuraminic acid containing
CC glycoproteins.
CC -I- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC -I- SUBCELLULAR LOCATION: External, anchored to the envelope by its N-
CC terminal hydrophobic sequence.
CC -I- SIMILARITY: Belongs to the paramyxoviruses hemagglutinin-
CC neuraminidase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y00114; CAA68298.1; -.
DR EMBL; D84095; BAA12218.1; -.
DR InterPro; IPR000665; Hem-neuramndse.
DR InterPro; IPR011040; Sialidase.
DR Pfam; PF00423; HN; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin; Hydrolase;
KW Transmembrane.
FT DOMAIN 1 30 Cytoplasmic (Potential) .
FT TRANSMEM 31 53 Potential.
FT DOMAIN 54 572 Extracellular (Potential) .
FT CARBOHYD 308 308 N-linked (GlcNAc... ) (Potential) .
FT CARBOHYD 351 351 N-linked (GlcNAc... ) (Potential) .
FT CARBOHYD 448 448 N-linked (GlcNAc... ) (Potential) .
FT CARBOHYD 523 523 N-linked (GlcNAc... ) (Potential) .
FT CARBOHYD 570 570 N-linked (GlcNAc... ) (Potential) .
SQ SEQUENCE 572 AA; 64590 MW; DDF84E90C4EC87B4 CRC64;

Query Match 70.6%; Score 36; DB 1; Length 572;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KASQDVGTSV 10
Db 82 KTSDDIGTSI 91

RESULT 13
Q65689 PRELIMINARY; PRT; 572 AA.
AC Q65689;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hemagglutinin-neuraminidase protein.
OS Bovine parainfluenza 3 virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
```

```
OC Paramyxoviridae; Paramyxovirinae; Respirovirus.
OX NCBI_TaxID=11215;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95287485; PubMed=7769692;
RA Breker-Klassen M.M., Yoo D., Mittal S.K., Sorden S.D., Haines D.M.,
RA Babiuk L.A.;
RT "Recombinant type 5 adenoviruses expressing bovine parainfluenza virus
RT type 3 glycoproteins protect Sigmodon hispidus cotton rats from bovine
RT parainfluenza virus type 3 infection."
RL J. Virol. 69:4308-4315(1995) .
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96405248; PubMed=8809388;
RA Breker-Klassen M.M., Yoo D., Babiuk L.A.;
RT "Comparisons of the F and HN gene sequences of different strains of
RT bovine parainfluenza virus type 3: relationship to phenotype and
RT pathogenicity."
RL Can. J. Vet. Res. 60:228-236(1996) .
RN [3]
RP SEQUENCE FROM N.A.
RA Breker-Klassen M.M.;
RT "The immunobiology of bovine parainfluenza virus type 3 F and HN
RT glycoproteins."
RL Thesis (1995), Veterinary Infectious Disease Organizatn, University of
RL Saskatchewan Unpublished.
CC -I- FUNCTION: Hemagglutinin is responsible for attaching the virus to
CC cell receptors and for initiating infection. Neuraminidase
CC activity helps the efficient spread of the virus by dissociating
CC the mature virions from the neuraminic acid containing
CC glycoproteins (by similarity) .
CC -I- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC -I- SUBCELLULAR LOCATION: External, anchored to the envelope by its N-
CC terminal hydrophobic sequence (by similarity) .
CC -I- SIMILARITY: Belongs to the paramyxoviruses hemagglutinin-
CC neuraminidase family.
CC EMBL; U31671; AAB03692.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0004308; F:exo-alpha-sialidase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0019058; P:viral infectious cycle; IEA.
DR InterPro; IPR000665; Hem-neuramndse.
DR InterPro; IPR011040; Sialidase.
DR Pfam; PF00423; HN; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin; Hydrolase;
KW Transmembrane.
SQ SEQUENCE 572 AA; 64668 MW; D708557A0CFA8B1C CRC64;

Query Match 70.6%; Score 36; DB 2; Length 572;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KASQDVGTSV 10
Db 82 KTSDDIGTSI 91

RESULT 14
CATA CAUCR STANDARD; PRT; 737 AA.
ID CATA CAUCR
AC O31066;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Peroxidase/catalase (EC 1.11.1.6) (Catalase-peroxidase) .
GN Name=katG; OrderedLocusNames=CC3043;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
```

```
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RA MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RX Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Uitterback T.R., Iran K., Wolf A.M., Vamathevan J.J.,
RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
RV [2]
RP SEQUENCE OF 1-494 FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=98012985; PubMed=9352936;
RA Steinman H.M., Fareed F., Weinstein L.;
RT "Catalase-peroxidase of Caulobacter crescentus: function and role in
stationary-phase survival.";
RL J. Bacteriol. 179:6831-6836(1997).
CC -!- FUNCTION: Bifunctional, exhibiting both a catalase and broad-
spectrum peroxidase activities.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: Binds 1 heme B (iron-protoporphyrin IX) group per dimer
(BY similarity).
CC -!- INDUCTION: By hydrogen peroxide.
CC -!- SIMILARITY: Belongs to the peroxidase family. Bacterial
peroxidase/catalase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005967; AAK25005.1; -.
DR EMBL; AF027168; AAC45850.1; -.
DR PIR; A87626; A87626.
DR PIR; T45480; T45480.
DR HSP; Q339D2; LMWV.
DR TIGR; CC3043; -.
DR InterPro; IPR000763; Bac_ctase/prase.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR010255; Peroxidase_super.
DR Pfam; PF00141; peroxidase; 1.
DR PRINTS; PR00460; PEROXIDASE.
DR PRINTS; PR00458; PEROXIDASE.
DR TIGRfam; TIGR00198; cat_per_HPI; 1.
DR PROSITE; PS00435; PEROXIDASE_1; 1.
DR PROSITE; PS00436; PEROXIDASE_2; 1.
DR PROSITE; PS50873; PEROXIDASE_4; 1.
KW Complete proteome; Heme; Hydrogen peroxide; Iron; Organic radical;
KW Oxidoreductase; Peroxidase.
FT ACT_SITE 102 102 Charge stabilization (By similarity).
FT ACT_SITE 106 106 Proton acceptor (By similarity).
FT ACT_SITE 319 319 Tryptophan radical intermediate (By
similarity).
FT METAL 268 268 Iron (heme axial ligand) (By similarity).
SQ SEQUENCE 737 AA; 80165 MW; D900742D04E02E97 CRC64;
```

Query Match 70.6%; Score 36; DB 1; Length 737;

Best Local Similarity 70.0%; Pred. No. 1.7e+02;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KASQDVGTSV 10

DB 550 KAAKDAGTSV 559

```
RESULT 15
Q7RNQ6 PRELIMINARY; PRT; 2598 AA.
AC Q7RNQ6;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PV01759;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=1236865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteza M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwylum T.V.,
RA Shallow S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Chao J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA Van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL; AABL01000473; ZAA21125.1; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR000225; Armadillo.
DR InterPro; IPR000362; Fumarate lyase.
DR PROSITE; PS0176; ARM_REPEAT; 1.
DR PROSITE; PS00163; FUMARATE_LYASES; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 2598 AA; 289931 MW; F2DEE98B4404D54E CRC64;

Query Match 70.6%; Score 36; DB 2; Length 2598;
Best Local Similarity 87.5%; Pred. No. 6.3e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 KASQDVGT 8

DB 2498 KASQDLGT 2505

Search completed: August 9, 2005, 14:19:35

Job time : 120.038 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: August 9, 2005, 14:02:28 ; Search time 81.8462 Seconds  
(without alignments)  
33.078 Million cell updates/sec

Title: US-10-089-452-28

Perfect score: 43

Sequence: 1 WTSTRHT 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	7	2	AAR97314 Humanised
2	43	100.0	7	4	AAB86085 H. pylori
3	43	100.0	7	4	AAB86053 H. pylori
4	43	100.0	7	8	ADM95103 Murine MN
5	43	100.0	7	8	ADM98312 Humanised
6	43	100.0	106	2	AAR97337 Humanised
7	43	100.0	106	2	AAR97335 Murine NE
8	43	100.0	106	4	AAB86108 H. pylori
9	43	100.0	106	4	AAB86070 H. pylori
10	43	100.0	106	8	ADM95097 Murine MN
11	43	100.0	106	8	ADM95100 Humanised
12	43	100.0	106	8	ADM98320 Murine MN
13	43	100.0	106	8	ADM98322 Humanised
14	43	100.0	108	7	ADD05266 Female mo
15	43	100.0	108	8	ADM80357 Murine ho
16	43	100.0	115	6	ABR42756 Anti-CEA
17	43	100.0	232	6	ABG76490 Light cha
18	43	100.0	232	7	ADE64204 MN14LC pr
19	43	100.0	232	7	ADF60817 hMN-14 li
20	43	100.0	250	2	AAW02278 741F8 ant
21	43	100.0	250	2	AAW29261 Anti-c-er
22	43	100.0	250	2	AAW22400 Single ch
23	43	100.0	250	2	AAW47012 Single ch
24	43	100.0	250	2	AAW53168 741F8 ant
25	43	100.0	250	2	AAW80422 Anti-c-er

RESULT 1  
AAR97314  
ID AAR97314 standard; peptide; 7 AA.

XX AAR97314;

XX 15-OCT-1996 (first entry)

DE Humanised monoclonal antibody light chain CDR.

XX Monoclonal antibody; humanised; mouse; framework region; FR; CDR;

KW complementary determining region; anti-carcinoembryonic antigen; CEA;

KW diagnosis; imaging; therapy; immune response.

XX Mus musculus.

XX WO9611013-A1.

XX 18-APR-1996.

XX 28-SEP-1995; 95WO-US011964.

XX 05-OCT-1994; 94US-00318157.

XX (IMMU-) IMMUNOMEDICS INC.

XX Hansen HJ, Armour KL;

XX WPI; 1996-209653/21.

XX New humanised anti-CEA monoclonal antibody - having engrafted murine CDRs, used for diagnosis, imaging and therapy of CEA-producing cancers.

XX Claim 5; Page 38; 62pp; English.

XX New humanised monoclonal antibodies (MAbs) comprising the complementary determining regions (CDRs) of a parental murine class III anti-carcinoembryonic (CEA) MAb engrafted to the framework regions (FR) of a heterologous antibody which can be derived from any species including human, retain the anti-CEA binding specificity of the parental murine MAb but are less immunogenic in a human subject than the parental MAb. The humanised antibodies can be used in diagnosis, imaging and therapy of CEA-producing cancers and patients receiving the humanised antibodies and conjugates show improved therapeutic results, decreased immune responses and decreased immune-mediated adverse effects compared to the parent antibody. This sequence is a murine derived CDR region which corresponds to CDR 2 in the light chain of the humanised MAb. See AAR97313-97333

Abw00714 741F8 sFV  
Abr42757 Anti-CEA  
Abr42758 Anti-CEA  
Abr42289 Diabody 6  
Abr42754 Anti-CEA  
Abr83645 hMN14-sCF  
Abr42295 Bispecifi  
Abr42293 Bispecifi  
Abr42298 Bispecifi  
Abr42290 Diabody h  
Abr42291 Bispecifi  
Abr83649 Tsl VL ch  
Abg76488 Humanised  
Abg74240 Chimaeric  
Aae27930 Human CSE  
Abb82839 Antibody  
Abu07879 Venezuela  
Abu07882 Venezuela  
Abu07880 Venezuela

## ALIGNMENTS

26	43	100.0	250	7	ABW00714
27	43	100.0	256	6	ABR42757
28	43	100.0	257	6	ABR42758
29	43	100.0	261	6	ABR42289
30	43	100.0	261	6	ABR42754
31	43	100.0	261	6	ABR83645
32	43	100.0	262	6	ABR42295
33	43	100.0	262	6	ABR42293
34	43	100.0	262	6	ABR42298
35	43	100.0	262	6	ABR42290
36	43	100.0	262	6	ABR42291
37	43	100.0	391	6	ABR83649
38	43	100.0	443	6	ABG76488
39	43	100.0	443	6	ABG74240
40	43	100.0	667	2	AAR39573
41	39	90.7	238	5	AAE27930
42	39	90.7	238	6	ABB82839
43	39	90.7	239	6	ABU07879
44	39	90.7	240	6	ABU07882
45	39	90.7	240	6	ABU07880

SQ Sequence 7 AA;  
 Query Match 100.0%; Score 43; DB 2; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WTSTRHT 7  
 |||||  
 Db 1 WTSTRHT 7

RESULT 3  
 AAB86053  
 ID AAB86053 standard; peptide; 7 AA.  
 XX  
 AC AAB86053;  
 XX  
 DT 17-JUL-2001 (first entry)  
 XX  
 DE H. pylori catalase derived antibody HP25/6m/1B5 light chain CDR2.  
 XX  
 KW Heavy chain; light chain; catalase; beta-urease; detection; infection; CDR; antigen;  
 KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
 KW antibacterial; complementarity determining region.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200127613-A2.  
 XX  
 PD 19-APR-2001.  
 XX  
 PF 12-OCT-2000; 2000WO-EP010058.  
 XX  
 PR 12-OCT-1999; 99EP-00120351.  
 PR 16-MAR-2000; 2000EP-00105592.  
 PR 31-MAR-2000; 2000EP-00107028.  
 PR 10-MAY-2000; 2000EP-00110110.  
 XX  
 PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
 XX  
 PI Reiter C, Cullmann G, Heppner P, Ringeis A, Mueller H, Haindl E;  
 XX  
 DR WPI: 2001-282087/29.  
 DR N-PSDB; AAF88055.  
 XX  
 PT Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
 PT sample.  
 XX  
 PS Claim 21; Page 16; 89pp; German.  
 XX  
 CC This invention describes a novel method for detecting, in a mammal,  
 CC infection by an acid-resistant microorganism (A) which comprises reacting  
 CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
 CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
 CC complex is formed with Ag, and the formation of a complex detected. R are  
 CC specific for an Ag which, after passage through the intestines, at least  
 CC in some mammals, retains a native (or corresponding) structure against  
 CC which the mammal produces antibodies (when immunized or infected with  
 CC (A), or its extracts, lysates or derived proteins (or fragments) or  
 CC synthetic peptides). The products of the invention have antibacterial  
 CC activity. The method is used to diagnose infection by Helicobacter.  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. Hepatica, C. Jejuni and M. tuberculosis, and also to monitor the  
 CC progress of treatment. Receptors, particularly antibodies, directed  
 CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an  
 CC infection. This sequence represents a complementarity determining region  
 CC (CDR) from an antibody generated against a Helicobacter pylori antigen  
 CC (catalase or beta-urease) which is used to illustrate the method of the  
 CC invention  
 XX  
 SQ Sequence 7 AA;  
 Query Match 100.0%; Score 43; DB 4; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WTSTRHT 7  
 |||||  
 Db 1 WTSTRHT 7

RESULT 2  
 AAB86085  
 ID AAB86085 standard; peptide; 7 AA.  
 XX  
 AC AAB86085;  
 XX  
 DT 17-JUL-2001 (first entry)  
 XX  
 DE H. pylori catalase derived antibody HP25/6m/1B5 light chain CDR2.  
 XX  
 KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;  
 KW acid-resistant microorganism; complementarity determining region; CDR;  
 KW feces; heavy chain; light chain.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200127613-A2.  
 XX  
 PD 19-APR-2001.  
 XX  
 PF 12-OCT-2000; 2000WO-EP010057.  
 XX  
 PR 12-OCT-1999; 99EP-00120351.  
 PR 16-MAR-2000; 2000EP-00105592.  
 PR 31-MAR-2000; 2000EP-00107028.  
 PR 10-MAY-2000; 2000EP-00110110.  
 XX  
 PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
 XX  
 PI Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;  
 XX  
 DR WPI: 2001-282086/29.  
 DR N-PSDB; AAF88112.  
 XX  
 PT Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing Helicobacter pylori, comprises immunochromatographic detection  
 PT of antigen in feces.  
 XX  
 PS Claim 26; Page 26; 90pp; German.  
 XX  
 CC This invention describes a novel method for detecting infection by an  
 CC acid-resistant microorganism (A), in a mammal, using  
 CC immunochromatography. The method is used to diagnose infection by an acid  
 CC -resistant microorganism (A), in a mammal, such as Helicobacter.  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. Hepatica, C. Jejuni and M. tuberculosis. The method is rapid, simple,  
 CC inexpensive and non-invasive, and may indicate the stage of infection. A  
 CC test strip used in the method may include a filter to eliminate particles  
 CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence represents a  
 CC complementarity determining region (CDR) from an antibody raised against  
 CC the H. pylori catalase or beta-urease antigen which is used to illustrate  
 CC the method of the invention  
 XX  
 SQ Sequence 7 AA;  
 Query Match 100.0%; Score 43; DB 4; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7  
 |||||

Db 1 WTSTRHT 7  
 |||||

RESULT 4  
 ADM95103  
 ID ADM95103 standard; peptide; 7 AA.  
 XX  
 AC ADM95103;  
 XX  
 XX  
 DT 15-JUL-2004 (first entry)  
 XX  
 DE Murine MN-14 MAB light chain variable region CDR2.  
 XX  
 KW non-medullary thyroid carcinoma; carcinoembryonic antigen; antibody;  
 KW complementarity determining region; pancreatic cancer; breast cancer;  
 KW ovarian cancer.  
 XX  
 XX Mus sp.  
 OS  
 XX WO2004032962-A1.  
 PN  
 XX 22-APR-2004.  
 PD  
 XX 11-OCT-2002; 2002WO-US032307.  
 PF  
 XX 08-OCT-2002; 2002US-0416531P.  
 PR  
 XX (IMMU-) IMMUNOMEDICS INC.  
 PA  
 XX Goldenberg DM, Hansen HJ;  
 PI  
 XX WPI; 2004-364813/34.  
 DR  
 XX  
 XX Treating non-medullary thyroid carcinoma or medullary thyroid carcinoma  
 PT comprises administering, either concurrently or sequentially, Class III  
 PT anti-CEA (carcinoembryonic antigen) monoclonal antibody or fragment and  
 PT therapeutic agent(s).  
 XX  
 XX Claim 6; SEQ ID NO 21; 89pp; English.  
 PS  
 XX The invention relates to a method of treating a non-medullary thyroid  
 CC carcinoma or treating medullary thyroid carcinoma involves administering  
 CC to a subject, either concurrently or sequentially, a Class III anti-CEA  
 CC (carcinoembryonic antigen) monoclonal antibody or its fragment and at  
 CC least one therapeutic agent. The antibody or its fragment comprises the  
 CC complementarity-determining regions (CDRs) of a murine MN-14 monoclonal  
 CC antibody. (I) is useful for treating medullary thyroid carcinoma or non-  
 CC medullary thyroid carcinoma such as pancreatic cancer, breast cancer,  
 CC ovarian cancer, in a subject. This sequence corresponds to the CDR2 of  
 CC the murine MN-14 antibody light chain variable region and used to  
 CC generate the humanised/chimeric antibody of the invention.  
 XX  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 43; DB 8; Length 7;  
 Best Local Similarity 100.0%; Pred. NO. 1.8e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7  
 |||||

Db 1 WTSTRHT 7  
 |||||

RESULT 5  
 ADM98312  
 ID ADM98312 standard; peptide; 7 AA.  
 XX  
 XX  
 AC ADM98312;  
 XX  
 DT 15-OCT-1996 (first entry)  
 XX

DT 15-JUL-2004 (first entry)  
 XX  
 DE Humanised murine MN-14 antibody light chain variable region CDR2 peptide.  
 XX  
 KW carcinoembryonic antigen; CEA; antibody therapy; anti-CEA; humanised;  
 KW murine; mouse; antibody; MN-14; non-medullary thyroid carcinoma;  
 KW medullary thyroid carcinoma; colorectal cancer; gastric cancer;  
 KW hepatocellular carcinoma; breast cancer; lung cancer;  
 KW antibody-dependent cell-mediated cytotoxicity; complement-mediated lysis;  
 KW apoptosis.  
 XX  
 OS Mus sp.  
 OS Synthetic.  
 PN WO2004032857-A2.  
 XX  
 PD 22-APR-2004.  
 XX  
 XX 08-OCT-2003; 2003WO-US031801.  
 PF  
 XX 08-OCT-2002; 2002US-0416531P.  
 PR  
 XX 11-OCT-2002; 2002WO-US032307.  
 PR  
 XX 02-MAY-2003; 2003US-0467161P.  
 XX  
 PA (IMMU-) IMMUNOMEDICS INC.  
 XX  
 XX Goldenberg DM, Hansen HJ;  
 PI  
 XX WPI; 2004-340797/31.  
 DR  
 XX Treating non-medullary or medullary thyroid carcinoma, involves  
 PT administering to subject, either concurrently or sequentially, anti-  
 PT carcinoembryonic antigen antibody or its fragment and therapeutic agent.  
 XX  
 XX Claim 8; Page 76; 122pp; English.  
 PS  
 XX This invention relates to methods for treating cancers that express  
 CC carcinoembryonic antigens (CEAs). Specifically, it refers to antibody  
 CC therapy and in particular anti-carcinoembryonic antigen (anti-CEA)  
 CC antibodies that can be used following administration of an agent that  
 CC will increase antibody effector functions or following administration of  
 CC an amount of interferon effective to upregulate CEA expression in tumour  
 CC cells. The present invention describes the anti-CEA antibody as a class  
 CC III humanised murine monoclonal antibody that retains specificity for CEA  
 CC binding and is preferably an MN-14 antibody or fragment thereof.  
 CC Accordingly, hMN-14 can be formulated into a cytostatic pharmaceutical  
 CC composition useful for the treatment of non-medullary or medullary  
 CC thyroid carcinoma, colorectal cancer, gastric cancer, hepatocellular  
 CC carcinoma, as well as breast or lung cancer. Furthermore, it can  
 CC stimulate antibody-dependent cell-mediated cytotoxicity, complement-  
 CC mediated lysis and apoptosis as well as inhibit cell cycle progression,  
 CC cell growth and angiogenesis. This peptide sequence is a CDR region of  
 CC the humanised murine MN-14 antibody of the invention.  
 XX  
 SQ Sequence 7 AA;

Query Match 100.0%; Score 43; DB 8; Length 7;  
 Best Local Similarity 100.0%; Pred. NO. 1.8e+06;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7  
 |||||

Db 1 WTSTRHT 7  
 |||||

RESULT 6  
 AAR97337  
 ID AAR97337 standard; protein; 106 AA.  
 XX  
 XX AAR97337;  
 AC  
 XX  
 DT 15-OCT-1996 (first entry)  
 XX

```

DE Humanised antibody variable light chain (MN14HuVK) .
XX
KW Monoclonal antibody; humanised; mouse; framework region; FR; CDR;
KW complementary determining region; anti-carcinoembryonic antigen; CEA;
KW diagnosis; imaging; therapy; immune response.
XX
OS Mus musculus.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Binding-site 24..34
FT Binding-site /label= CDR 1.
FT Binding-site 50..56
FT Binding-site /label= CDR 2.
FT Binding-site 89..96
FT Binding-site /label= CDR 3.
XX
XX WO9611013-A1.
XX
XX 18-APR-1996.
XX
XX 28-SEP-1995; 95WO-US011964.
XX
XX 05-OCT-1994; 94US-00318157.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Hansen HJ, Armour KL;
XX
XX WPI; 1996-209653/21.
XX
XX N-PSDB; AAT29013.
XX
XX New humanised anti-CEA monoclonal antibody - having engrafted murine
XX CDRs, used for diagnosis, imaging and therapy of CEA-producing cancers.
XX
XX Disclosure; Fig 8; 62pp; English.
XX
XX New humanised monoclonal antibodies (MAbs) comprising the complementary
XX determining regions (CDRs) of a parental murine class III anti-
XX carcinoembryonic (CEA) MAb engrafted to the framework regions (FRs) of a
XX heterologous antibody which can be derived from any species including
XX human, retain the anti-CEA binding specificity of the parental murine MAb
XX but are less immunogenic in a human subject than the parental MAb. The
XX humanised antibodies can be used in diagnosis, imaging and therapy of CEA
XX -producing cancers and patients receiving the humanised antibodies and
XX conjugates show improved therapeutic results, decreased immune responses
XX and decreased immune-mediated adverse effects compared to the parent
XX antibody. See AAR97313-97333 for the CDR sequence and framework regions
XX of the humanised MAbs
XX
XX Sequence 106 AA;
XX
XX Query Match 100.0%; Score 43; DB 2; Length 106;
XX Best Local Similarity 100.0%; Pred. No. 2.1;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7
DB 50 WTSTRHT 56

RESULT 7
AAR97335
ID AAR97335 standard; protein; 106 AA.
XX
XX AAR97335;
XX
XX 15-OCT-1996 (first entry)
XX
XX Murine NEMM MN-14 variable region light chain.
XX
XX Monoclonal antibody; humanised; mouse; framework region; FR; CDR;
KW complementary determining region; anti-carcinoembryonic antigen; CEA;

diagnosis; imaging; therapy; immune response.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Binding-site 24..34
FT Binding-site /label= CDR 1.
FT Binding-site 50..56
FT Binding-site /label= CDR 2.
FT Binding-site 89..96
FT Binding-site /label= CDR 3.
XX
XX WO9611013-A1.
XX
XX 18-APR-1996.
XX
XX 28-SEP-1995; 95WO-US011964.
XX
XX 05-OCT-1994; 94US-00318157.
XX
XX (IMMU-) IMMUNOMEDICS INC.
XX
XX Hansen HJ, Armour KL;
XX
XX WPI; 1996-209653/21.
XX
XX N-PSDB; AAT29011.
XX
XX New humanised anti-CEA monoclonal antibody - having engrafted murine
XX CDRs, used for diagnosis, imaging and therapy of CEA-producing cancers.
XX
XX Example 7; Fig 2; 62pp; English.
XX
XX New humanised monoclonal antibodies (MAbs) comprising the complementary
XX determining regions (CDRs) of a parental murine class III anti-
XX carcinoembryonic (CEA) MAb engrafted to the framework regions (FRs) of a
XX heterologous antibody which can be derived from any species including
XX human, retain the anti-CEA binding specificity of the parental murine MAb
XX but are less immunogenic in a human subject than the parental MAb. The
XX humanised antibodies can be used in diagnosis, imaging and therapy of CEA
XX -producing cancers and patients receiving the humanised antibodies and
XX conjugates show improved therapeutic results, decreased immune responses
XX and decreased immune-mediated adverse effects compared to the parent
XX antibody. See AAR97313-97333 for the CDR sequence and framework regions
XX of the humanised MAbs
XX
XX Sequence 106 AA;
XX
XX Query Match 100.0%; Score 43; DB 2; Length 106;
XX Best Local Similarity 100.0%; Pred. No. 2.1;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7
DB 50 WTSTRHT 56

RESULT 8
AAR86108
ID AAR86108 standard; protein; 106 AA.
XX
XX AAR86108;
XX
XX 17-JUL-2001 (first entry)
XX
XX H. pylori catalase antibody HP25/2m/1BS V region light chain.
XX
XX Catalase; beta-urease; antibody; antigen; detection; infection; epitope;
KW acid-resistant microorganism; complementarity determining region; CDR;
KW feces; heavy chain; light chain.
XX
XX Unidentified.
XX
XX WO200127612-A2.

```

[illegible]

XX (IMMU-) IMMUNOMEDICS INC.  
 XX Goldenberg DM, Hansen HJ;  
 XX WPI; 2004-364813/34.  
 DR N-PSDB; ADM95096.  
 XX Treating non-medullary thyroid carcinoma or medullary thyroid carcinoma  
 PT comprises administering, either concurrently or sequentially, Class III  
 PT anti-CEA (carcinoembryonic antigen) monoclonal antibody or fragment and  
 PT therapeutic agent (s).  
 XX Disclosure; Fig 12A-B; 88pp; English.  
 PS The invention relates to a method of treating a non-medullary thyroid  
 XX carcinoma or treating medullary thyroid carcinoma involves administering  
 CC to a subject, either concurrently or sequentially, a Class III anti-CEA  
 CC (carcinoembryonic antigen) monoclonal antibody or its fragment and at  
 CC least one therapeutic agent. The antibody or its fragment comprises the  
 CC complementarity-determining regions (CDRs) of a murine MN-14 monoclonal  
 CC antibody. (I) is useful for treating medullary thyroid carcinoma or non-  
 CC medullary thyroid carcinoma such as pancreatic cancer, breast cancer,  
 CC ovarian cancer, in a subject. This sequence corresponds to the murine MN-  
 CC 14 antibody light chain variable region.  
 XX Sequence 106 AA;  
 SQ

Query Match 100.0%; Score 43; DB 8; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 2.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
 DB 50 WTSTRHT 56  
 |||||||

RESULT 11  
 ADM95100  
 ID ADM95100 standard; protein; 106 AA.  
 XX  
 AC ADM95100;  
 XX  
 DT 15-JUL-2004 (first entry)  
 XX  
 DE Humanised murine MN-14 MAB light chain variable region.  
 XX  
 KW non-medullary thyroid carcinoma; carcinoembryonic antigen; antibody;  
 KW complementarity determining region; pancreatic cancer; breast cancer;  
 KW ovarian cancer.  
 XX  
 OS Mus sp.  
 OS Homo sapiens.  
 XX  
 PN WO2004032962-A1.  
 XX  
 PD 22-APR-2004.  
 XX  
 PF 11-OCT-2002; 2002WO-US032307.  
 XX  
 PR 08-OCT-2002; 2002US-0416531P.  
 XX  
 PA (IMMU-) IMMUNOMEDICS INC.  
 XX  
 PI Goldenberg DM, Hansen HJ;  
 XX  
 DR WPI; 2004-364813/34.  
 DR N-PSDB; ADM95096.  
 XX Treating non-medullary thyroid carcinoma or medullary thyroid carcinoma  
 PT comprises administering, either concurrently or sequentially, Class III  
 PT anti-CEA (carcinoembryonic antigen) monoclonal antibody or fragment and  
 PT therapeutic agent (s).

XX Claim 14; Fig 16A; 88pp; English.  
 XX The invention relates to a method of treating a non-medullary thyroid  
 CC carcinoma or treating medullary thyroid carcinoma involves administering  
 CC to a subject, either concurrently or sequentially, a Class III anti-CEA  
 CC (carcinoembryonic antigen) monoclonal antibody or its fragment and at  
 CC least one therapeutic agent. The antibody or its fragment comprises the  
 CC complementarity-determining regions (CDRs) of a murine MN-14 monoclonal  
 CC antibody. (I) is useful for treating medullary thyroid carcinoma or non-  
 CC medullary thyroid carcinoma such as pancreatic cancer, breast cancer,  
 CC ovarian cancer, in a subject. This sequence corresponds to the humanised  
 CC murine MN-14 antibody light chain variable region.  
 XX Sequence 106 AA;  
 SQ

Query Match 100.0%; Score 43; DB 8; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 2.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
 DB 50 WTSTRHT 56  
 |||||||

RESULT 12  
 ADM98320  
 ID ADM98320 standard; protein; 106 AA.  
 XX  
 AC ADM98320;  
 XX  
 DT 15-JUL-2004 (first entry)  
 XX  
 DE Murine MN-14 antibody light chain variable region protein.  
 XX  
 KW carcinoembryonic antigen; CEA; antibody therapy; anti-CEA; murine; mouse;  
 KW antibody; MN-14; non-medullary thyroid carcinoma;  
 KW medullary thyroid carcinoma; colorectal cancer; gastric cancer;  
 KW hepatocellular carcinoma; breast cancer; lung cancer;  
 KW antibody-dependent cell-mediated cytotoxicity; complement-mediated lysis;  
 KW apoptosis.  
 XX  
 OS Mus sp.  
 OS  
 PN WO2004032857-A2.  
 XX  
 PD 22-APR-2004.  
 XX  
 PF 08-OCT-2003; 2003WO-US031801.  
 XX  
 PR 08-OCT-2002; 2002US-0416531P.  
 PR 11-OCT-2002; 2002WO-US032307.  
 PR 02-MAY-2003; 2003US-0467161P.  
 XX  
 PA (IMMU-) IMMUNOMEDICS INC.  
 XX  
 PI Goldenberg DM, Hansen HJ;  
 XX  
 DR WPI; 2004-340797/31.  
 DR N-PSDB; ADM98319.  
 XX  
 PT Treating non-medullary or medullary thyroid carcinoma, involves  
 PT administering to subject, either concurrently or sequentially, anti-  
 PT carcinoembryonic antigen antibody or its fragment and therapeutic agent.  
 XX  
 PS Disclosure; Fig 12; 122pp; English.  
 XX  
 CC This invention relates to methods for treating cancers that express  
 CC carcinoembryonic antigens (CEAs). Specifically, it refers to antibody  
 CC therapy and in particular anti-carcinoembryonic antigen (anti-CEA)  
 CC antibodies that can be used following administration of an agent that  
 CC will increase antibody effector functions or following administration of  
 CC an amount of interferon effective to upregulate CEA expression in tumour

CC cells. The present invention describes the anti-CEA antibody as a class  
 CC III humanised murine monoclonal antibody that retains specificity for CEA  
 CC binding and is preferably an MN-14 antibody or fragment thereof.  
 CC Accordingly, hMN-14 can be formulated into a cytostatic pharmaceutical  
 CC composition useful for the treatment of non-medullary or medullary  
 CC thyroid carcinoma, colorectal cancer, gastric cancer, hepatocellular  
 CC carcinoma, as well as breast or lung cancer. Furthermore, it can  
 CC stimulate antibody-dependent cell-mediated cytotoxicity, complement-  
 CC mediated lysis and apoptosis as well as inhibit cell cycle progression,  
 CC cell growth and angiogenesis. This polypeptide sequence is murine MN-14  
 CC antibody light chain variable region protein of the invention.  
 XX  
 XX Sequence 106 AA;

Query Match 100.0%; Score 43; DB 8; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 2.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
 |||||  
 Db 50 WTSTRHT 56

RESULT 13  
 ADM98322  
 ID ADM98322 standard; protein; 106 AA.  
 XX  
 AC ADM98322;  
 DT 15-JUL-2004 (first entry)  
 XX  
 DE Humanised murine MN-14 antibody light chain variable region protein.  
 XX  
 KW carcinoembryonic antigen; CEA; antibody therapy; anti-CEA; murine; mouse;  
 KW antibody; MN-14; non-medullary thyroid carcinoma;  
 KW medullary thyroid carcinoma; colorectal cancer; gastric cancer;  
 KW hepatocellular carcinoma; breast cancer; lung cancer;  
 KW antibody-dependent cell-mediated cytotoxicity; complement-mediated lysis;  
 KW apoptosis; humanised.  
 XX  
 OS Mus sp.  
 OS Synthetic.  
 XX  
 PN WO2004032857-A2.  
 XX  
 PD 22-APR-2004.  
 XX  
 PF 08-OCT-2003; 2003WO-US031801.  
 XX  
 PR 08-OCT-2002; 2002US-0416531P.  
 PR 11-OCT-2002; 2002WO-US032307.  
 PR 02-MAY-2003; 2003US-0467161P.  
 XX  
 PA (IMMU-) IMMUNOMEDICS INC.  
 XX  
 PI Goldenberg DM, Hansen HJ;  
 XX  
 DR WPI; 2004-340797/31.  
 DR N-PSDB; ADM98321.  
 XX  
 PT Treating non-medullary or medullary thyroid carcinoma, involves  
 PT administering to subject, either concurrently or sequentially, anti-  
 PT carcinoembryonic antigen antibody or its fragment and therapeutic agent.  
 XX  
 PS Disclosure; Fig 23a; 122pp; English.  
 XX  
 CC This invention relates to methods for treating cancers that express  
 CC carcinoembryonic antigens (CEAs). Specifically, it refers to antibody  
 CC therapy and in particular anti-carcinoembryonic antigen (anti-CEA)  
 CC antibodies that can be used following administration of an agent that  
 CC will increase antibody effector functions or following administration of  
 CC an amount of interferon effective to upregulate CEA expression in tumour  
 CC cells. The present invention describes the anti-CEA antibody as a class

CC III humanised murine monoclonal antibody that retains specificity for CEA  
 CC binding and is preferably an MN-14 antibody or fragment thereof.  
 CC Accordingly, hMN-14 can be formulated into a cytostatic pharmaceutical  
 CC composition useful for the treatment of non-medullary or medullary  
 CC thyroid carcinoma, colorectal cancer, gastric cancer, hepatocellular  
 CC carcinoma, as well as breast or lung cancer. Furthermore, it can  
 CC stimulate antibody-dependent cell-mediated cytotoxicity, complement-  
 CC mediated lysis and apoptosis as well as inhibit cell cycle progression,  
 CC cell growth and angiogenesis. This polypeptide sequence is the humanised  
 CC murine MN-14 antibody light chain variable region protein of the  
 CC invention.  
 XX  
 XX Sequence 106 AA;

Query Match 100.0%; Score 43; DB 8; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 2.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
 |||||  
 Db 50 WTSTRHT 56

RESULT 14  
 ADD05266  
 ID ADD05266 standard; protein; 108 AA.  
 XX  
 AC ADD05266;  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE Female mouse sex hormone of the invention #2.  
 XX  
 KW female sex hormone.  
 XX  
 OS Mus sp.  
 PN WO2003074704-A1.  
 XX  
 PD 12-SEP-2003.  
 XX  
 PF 28-FEB-2003; 2003WO-JP002311.  
 XX  
 PR 01-MAR-2002; 2002JP-00055669.  
 XX  
 PA (TAXE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Katagiri M, Fujimoto S, Goda Y;  
 XX  
 DR WPI; 2003-731681/69.  
 DR N-PSDB; ADD05265.  
 XX  
 PT Novel proteins for binding, identifying and concentrating female sex  
 PT hormones.  
 XX  
 PS Claim 1; SEQ ID NO 4; 101pp; Japanese.  
 XX  
 CC The present invention relates to proteins that bind to female sex  
 CC hormones. The method is useful for binding, identifying and concentrating  
 CC female sex hormones. The present invention represents a female sex  
 CC hormone of the invention.  
 XX  
 XX Sequence 108 AA;

Query Match 100.0%; Score 43; DB 7; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 2.2;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
 |||||  
 Db 50 WTSTRHT 56

RESULT 15  
ADM80357  
ID ADM80357 standard; protein; 108 AA.  
XX  
AC ADM80357;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Murine hormone disruptor E2-73 antibody (scFv) protein SeqID 4.  
XX  
KW antibody; mouse; murine; scFv; hormone disruptor; alkylphenol;  
KW resin component; chlorophenol; immunological isolation;  
KW environmental protection.  
XX  
OS Mus sp.  
XX  
FN WO2003101611-A1.  
XX  
PD 11-DEC-2003.  
XX  
PE 30-MAY-2003; 2003WO-JP006840.  
XX  
PR 31-MAY-2002; 2002JP-00159728.  
XX  
PA (NIEN-) JAPAN ENVIROCHEMICALS LTD.  
PA (ENVI-) ENVIRONMENT PURIFICATION RES INST INC.  
XX  
FI Saito K, Shiraishi T, Goda Y;  
XX  
DR WPI; 2004-043011/04.  
DR N-PSDB; ADM80356.  
XX  
PT Polymer-molded support carrying anti-hormone-disrupting-substance  
PT antibody for immunological isolation, concentration or clean-up of such  
PT trace hormone disruptors in samples, applicable in environmental  
PT protection.  
XX  
PS Example 7; SEQ ID NO 4; 133pp; Japanese.  
XX  
CC This invention relates to a novel material useful for the isolation of  
CC hormone disrupting substances (i.e. hormone disruptors). Specifically, it  
CC refers to a material that comprises a polymer-molded support carrying an  
CC anti-hormone disruptor antibody covalently linked via a graft chain to  
CC the support, which can be used to capture and subsequently recover the  
CC substances of interest. The present invention provides a method for  
CC selectively and efficiently concentrating hormone disruptors including  
CC alkylphenols, resin components, chlorophenols and female, male or thyroid  
CC hormones from environmental samples. In particular, the method is useful  
CC for the immunological isolation, concentration or clean-up of such trace  
CC hormone disruptors that occur in river water, such that it is a method  
CC applicable for environmental protection and the analytical sciences. This  
CC polypeptide is a murine hormone disruptor antibody (scFv) protein  
CC sequence of the invention.  
XX  
SQ Sequence 108 AA;  
Query Match 100.0%; Score 43; DB 8; Length 108;  
Best Local Similarity 100.0%; Pred. No. 2.2; 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0;  
Qy 1 WTSTRHT 7  
| | | | |  
Db 50 WTSTRHT 56  
Search completed: August 9, 2005, 14:14:49  
Job time : 82.8462 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:07:04 ; Search time 21 Seconds  
(without alignments)  
24.883 Million cell updates/sec

Title: US-10-089-452-28  
Perfect score: 43  
Sequence: 1 WTSTRHT 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCITUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	7	2	US-08-318-157B-21
2	43	100.0	7	4	US-09-253-794-21
3	43	100.0	106	2	US-08-318-157B-4
4	43	100.0	106	2	US-08-318-157B-19
5	43	100.0	106	4	US-09-253-794-4
6	43	100.0	106	4	US-09-253-794-19
7	43	100.0	250	1	US-08-133-804-2
8	43	100.0	250	1	US-08-461-184-8
9	43	100.0	250	1	US-08-463-675-8
10	43	100.0	250	1	US-08-464-589-8
11	43	100.0	250	1	US-08-461-838-2
12	43	100.0	250	2	US-08-461-386-2
13	43	100.0	622	2	US-08-356-786-16
14	39	90.7	110	4	US-10-092-246-13
15	39	90.7	110	4	US-10-096-246A-13
16	39	90.7	239	4	US-10-092-246-32
17	39	90.7	239	4	US-10-092-246-33
18	39	90.7	239	4	US-10-096-246A-32
19	39	90.7	239	4	US-10-096-246A-33
20	39	90.7	240	4	US-10-092-246-34
21	39	90.7	240	4	US-10-092-246-35
22	39	90.7	240	4	US-10-092-246-36
23	39	90.7	240	4	US-10-092-246-37
24	39	90.7	240	4	US-10-096-246A-34
25	39	90.7	240	4	US-10-096-246A-35
26	39	90.7	240	4	US-10-096-246A-36
27	39	90.7	240	4	US-10-096-246A-37

28	39	90.7	257	4	US-10-092-246-2	Sequence 2, Appli
29	39	90.7	257	4	US-10-096-246A-2	Sequence 2, Appli
30	38	88.4	106	3	US-09-189-129-4	Sequence 4, Appli
31	38	88.4	106	4	US-09-824-286-4	Sequence 4, Appli
32	38	88.4	107	1	US-07-634-278-50	Sequence 50, Appli
33	38	88.4	107	1	US-07-634-278-51	Sequence 51, Appli
34	38	88.4	107	1	US-08-477-728-50	Sequence 50, Appli
35	38	88.4	107	1	US-08-477-728-51	Sequence 51, Appli
36	38	88.4	107	1	US-08-474-040-50	Sequence 50, Appli
37	38	88.4	107	1	US-08-474-040-51	Sequence 51, Appli
38	38	88.4	107	1	US-08-487-200-50	Sequence 50, Appli
39	38	88.4	107	1	US-08-487-200-51	Sequence 51, Appli
40	38	88.4	107	1	US-08-491-845-4	Sequence 4, Appli
41	38	88.4	107	1	US-08-491-845-12	Sequence 12, Appli
42	38	88.4	107	3	US-08-838-682-19	Sequence 19, Appli
43	38	88.4	107	3	US-08-895-914-19	Sequence 19, Appli
44	38	88.4	107	3	US-08-484-537-50	Sequence 50, Appli
45	38	88.4	107	3	US-08-484-537-51	Sequence 51, Appli

ALIGNMENTS

RESULT 1  
US-08-318-157B-21  
; Sequence 21, Application US/08318157B  
; Patent No. 5874540  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; APPLICANT: ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/318,157B  
; FILING DATE: 05-OCT-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665  
; REFERENCE/DOCKET NUMBER: 18733/464  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: amino acid  
; LENGTH: 7 amino acids  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-318-157B-21

Query Match 100.0%; Score 43; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 WTSTRHT 7  
Db 1 WTSTRHT 7

```
RESULT 2
US-09-253-794-21
; Sequence 21, Application US/09253794
; Patent No. 6676924
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/253,794
; FILING DATE: 22-Feb-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,157
; FILING DATE: 05-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-253-794-21
Query Match 100.0%; Score 43; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7
Db 1 WTSTRHT 7

RESULT 3
US-08-318-157B-4
; Sequence 4, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
```

```
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-318-157B-4
Query Match 100.0%; Score 43; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7
Db 50 WTSTRHT 56

RESULT 4
US-08-318-157B-19
; Sequence 19, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
```

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-318-157B-19
Query Match 100.0%; Score 43; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7
Db 50 WTSTRHT 56

RESULT 5
US-09-253-794-4
; Sequence 4, Application US/09253794
; Patent No. 6676924
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/253,794
; FILING DATE: 22-Feb-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,157
; FILING DATE: 05-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-253-794-4
Query Match 100.0%; Score 43; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7
Db 50 WTSTRHT 56

RESULT 6
US-09-253-794-19
; Sequence 19, Application US/09253794
; Patent No. 6676924
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/253,794
; FILING DATE: 22-Feb-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,157
; FILING DATE: 05-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-253-794-4
Query Match 100.0%; Score 43; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7
Db 50 WTSTRHT 56

RESULT 7
US-08-133-804-2
; Sequence 2, Application US/08133804
; Patent No. 5534254
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/133,804
/ FILING DATE:
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kelley, Robin D.
/ REGISTRATION NUMBER: 34,637
/ REFERENCE/DOCKET NUMBER: 2054/22
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-248-7477
/ TELEFAX: 617-248-7100
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 250 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-133-804-2

Query Match      100.0%; Score 43; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WTSTRHT 7
DB      185 WTSTRHT 191

RESULT 8
US-08-461-184-8
/ Sequence 8, Application US/08461184
/ Patent No. 5631159
/ GENERAL INFORMATION:
/ APPLICANT: DORAI, HAIMANTI
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN
/ TITLE OF INVENTION: PRODUCTION FROM NON-NATIVE DNA
/ NUMBER OF SEQUENCES: 10
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC
/ STREET: 45 SOUTH STREET
/ CITY: HOPKINTON
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 07148
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/461,184
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION NUMBER: US/08/143,498
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KELLEY, ROBIN D
/ REGISTRATION NUMBER: 34,637
/ REFERENCE/DOCKET NUMBER: CRP093
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/248-7000
/ TELEFAX: 617/248-7100
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 250 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-461-184-8
```

```
Query Match      100.0%; Score 43; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WTSTRHT 7
DB      185 WTSTRHT 191

RESULT 9
US-08-463-675-8
/ Sequence 8, Application US/08463675
/ Patent No. 5658763
/ GENERAL INFORMATION:
/ APPLICANT: DORAI, HAIMANTI
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN
/ TITLE OF INVENTION: PRODUCTION FROM NON-NATIVE DNA
/ NUMBER OF SEQUENCES: 10
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC
/ STREET: 45 SOUTH STREET
/ CITY: HOPKINTON
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 07148
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/463,675
/ FILING DATE: 05-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/143,498
/ FILING DATE: 25-OCT-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KELLEY, ROBIN D
/ REGISTRATION NUMBER: 34,637
/ REFERENCE/DOCKET NUMBER: CRP093
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/248-7000
/ TELEFAX: 617/248-7100
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 250 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-463-675-8

Query Match      100.0%; Score 43; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WTSTRHT 7
DB      185 WTSTRHT 191

RESULT 10
US-08-464-589-8
/ Sequence 8, Application US/08464589
/ Patent No. 5733782
/ GENERAL INFORMATION:
/ APPLICANT: DORAI, HAIMANTI
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN
/ TITLE OF INVENTION: PRODUCTION FROM NON-NATIVE DNA
/ NUMBER OF SEQUENCES: 10
```

;/ CORRESPONDENCE ADDRESS:  
;/ ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC  
;/ STREET: 45 SOUTH STREET  
;/ CITY: HOPKINTON  
;/ STATE: MA  
;/ COUNTRY: USA  
;/ ZIP: 07148  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Floppy disk  
;/ COMPUTER: IBM PC compatible  
;/ OPERATING SYSTEM: PC-DOS/MS-DOS  
;/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/08/464,589  
;/ FILING DATE: 05-JUN-1995  
;/ CLASSIFICATION: 435  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: US 08/143,498  
;/ FILING DATE: 25-OCT-1993  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: KELLEY, ROBIN D  
;/ REGISTRATION NUMBER: 34,637  
;/ REFERENCE/DOCKET NUMBER: CRP093  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: 617/248-7000  
;/ TELEFAX: 617/248-7100  
;/ INFORMATION FOR SEQ ID NO: 8:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 250 amino acids  
;/ TYPE: amino acid  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: protein  
;/ US-08-464-589-8

Query Match 100.0%; Score 43; DB 1; Length 250;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7  
Db 185 WTSTRHT 191

RESULT 11  
US-08-461-838-2  
;/ Sequence 2, Application US/08461838  
;/ Patent No. 5753204  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Huston, James S.  
;/ APPLICANT: Oppermann, Hermann  
;/ APPLICANT: Houston, L. L.  
;/ APPLICANT: Ring, David B.  
;/ TITLE OF INVENTION: Biosynthetic Binding Proteins For  
;/ NUMBER OF SEQUENCES: 11  
;/ CORRESPONDENCE ADDRESS:  
;/ ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department  
;/ STREET: Exchange Place, 53 State Street  
;/ CITY: Boston  
;/ STATE: Massachusetts  
;/ COUNTRY: USA  
;/ ZIP: 02109  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Floppy disk  
;/ COMPUTER: IBM PC compatible  
;/ OPERATING SYSTEM: PC-DOS/MS-DOS  
;/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/08/461,838  
;/ FILING DATE:  
;/ CLASSIFICATION: 424  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Kelley, Robin D.

;/ REGISTRATION NUMBER: 34,637  
;/ REFERENCE/DOCKET NUMBER: 2054/22  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: 617-248-7477  
;/ TELEFAX: 617-248-7100  
;/ INFORMATION FOR SEQ ID NO: 2:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 250 amino acids  
;/ TYPE: amino acid  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: protein  
;/ US-08-461-838-2

Query Match 100.0%; Score 43; DB 1; Length 250;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7  
Db 185 WTSTRHT 191

RESULT 12  
US-08-461-386-2  
;/ Sequence 2, Application US/08461386  
;/ Patent No. 5837846  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Huston, James S.  
;/ APPLICANT: Oppermann, Hermann  
;/ APPLICANT: Houston, L. L.  
;/ APPLICANT: Ring, David B.  
;/ TITLE OF INVENTION: Biosynthetic Binding Proteins For  
;/ NUMBER OF SEQUENCES: 11  
;/ CORRESPONDENCE ADDRESS:  
;/ ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department  
;/ STREET: Exchange Place, 53 State Street  
;/ CITY: Boston  
;/ STATE: Massachusetts  
;/ COUNTRY: USA  
;/ ZIP: 02109  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Floppy disk  
;/ COMPUTER: IBM PC compatible  
;/ OPERATING SYSTEM: PC-DOS/MS-DOS  
;/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/08/461,386  
;/ FILING DATE:  
;/ CLASSIFICATION: 424  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Kelley, Robin D.  
;/ REGISTRATION NUMBER: 34,637  
;/ REFERENCE/DOCKET NUMBER: 2054/22  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: 617-248-7477  
;/ TELEFAX: 617-248-7100  
;/ INFORMATION FOR SEQ ID NO: 2:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 250 amino acids  
;/ TYPE: amino acid  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: protein  
;/ US-08-461-386-2

Query Match 100.0%; Score 43; DB 2; Length 250;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7  
Db 185 WTSTRHT 191

RESULT 13  
US-08-356-786-16  
; Sequence 16, Application US/08356786  
; GENERAL INFORMATION:  
; Patent No. 5877305  
; APPLICANT: Huston, James S.  
; APPLICANT: Oppermann, Hermann  
; APPLICANT: Houston, L. L.  
; APPLICANT: Ring, David B.  
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
; TITLE OF INVENTION: Marker  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
; STREET: Exchange Place, 53 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/356,786  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/831,967  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitcher, Edmund R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: CRP-053  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 622 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: Protein  
US-08-356-786-16

Query Match 100.0%; Score 43; DB 2; Length 622;  
Best Local Similarity 100.0%; Pred. No. 6.7; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0;

QY 1 WTSTRHT 7  
Db 185 WTSTRHT 191

RESULT 14  
US-10-092-246-13  
; Sequence 13, Application US/10092246  
; Patent No. 6501314  
; GENERAL INFORMATION:  
; APPLICANT: The Minister of National Defence, Government of Canada  
; APPLICANT: Fulton, R E  
; APPLICANT: Alvi, Azhar E  
; APPLICANT: Nagata, Leslie  
; TITLE OF INVENTION: Cloning, Expression, Sequencing, and Functional Enhancement of Mo  
; TITLE OF INVENTION: scFv Antibody Against Venezuelan Equine Encephalitis Virus (Vee)  
; FILE REFERENCE: NEL-0007  
; CURRENT APPLICATION NUMBER: US/10/092,246  
; CURRENT FILING DATE: 2002-03-13  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13

; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Mouse hybridoma cell line 1A4A1  
US-10-092-246-13

Query Match 90.7%; Score 39; DB 4; Length 110;  
Best Local Similarity 85.7%; Pred. No. 6.1;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
Db 49 WSSTRHT 55

RESULT 15  
US-10-096-246A-13  
; Sequence 13, Application US/10096246A  
; Patent No. 6818748  
; GENERAL INFORMATION:  
; APPLICANT: Fulton, R. Elaine  
; APPLICANT: Nagata, Leslie  
; APPLICANT: Alvi, Azhar Z.  
; TITLE OF INVENTION: Cloning, Expression, Sequencing, and Functional Enhancement of  
; TITLE OF INVENTION: Monoclonal ScFv Antibody Against Venezuelan Equine Encephalitis  
; FILE REFERENCE: NEL-0007  
; CURRENT APPLICATION NUMBER: US/10/096,246A  
; CURRENT FILING DATE: 2002-03-13  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 13  
; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: mouse  
US-10-096-246A-13

Query Match 90.7%; Score 39; DB 4; Length 110;  
Best Local Similarity 85.7%; Pred. No. 6.1;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
Db 49 WSSTRHT 55

Search completed: August 9, 2005, 14:22:06  
Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:09:49 ; Search time 75.6538 Seconds  
(without alignments)  
36.122 Million cell updates/sec

Title: US-10-089-452-28  
Perfect score: 43  
Sequence: 1 WTSTRHT 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1752860 seqs, 390397842 residues

Total number of hits satisfying chosen parameters: 1752860

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/FCI\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/FCIUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	7	9	US-09-253-794-21
2	43	100.0	7	16	US-10-680-734-21
3	43	100.0	7	17	US-10-755-382-21
4	43	100.0	106	9	US-09-253-794-4
5	43	100.0	106	9	US-09-253-794-19
6	43	100.0	106	16	US-10-680-734-4
7	43	100.0	106	16	US-10-680-734-19
8	43	100.0	106	17	US-10-755-382-4
9	43	100.0	106	17	US-10-755-382-19
10	43	100.0	115	14	US-10-270-073-4
11	43	100.0	232	13	US-10-006-771A-6

12	43	100.0	232	15	US-10-377-109-4	Sequence 4, Appli
13	43	100.0	250	9	US-09-887-853-2	Sequence 2, Appli
14	43	100.0	250	17	US-10-683-547-2	Sequence 2, Appli
15	43	100.0	256	14	US-10-270-073-6	Sequence 6, Appli
16	43	100.0	257	14	US-10-270-073-8	Sequence 8, Appli
17	43	100.0	261	14	US-10-270-071-18	Sequence 18, Appli
18	43	100.0	261	14	US-10-270-073-2	Sequence 2, Appli
19	43	100.0	261	14	US-10-328-150-6	Sequence 6, Appli
20	43	100.0	262	14	US-10-270-071-20	Sequence 20, Appli
21	43	100.0	262	14	US-10-270-071-24	Sequence 24, Appli
22	43	100.0	262	14	US-10-270-071-28	Sequence 28, Appli
23	43	100.0	262	14	US-10-270-071-34	Sequence 34, Appli
24	43	100.0	358	17	US-10-829-388-12	Sequence 12, Appli
25	43	100.0	363	17	US-10-829-388-2	Sequence 2, Appli
26	43	100.0	364	17	US-10-829-388-11	Sequence 11, Appli
27	43	100.0	370	17	US-10-829-388-1	Sequence 1, Appli
28	43	100.0	391	14	US-10-328-150-14	Sequence 14, Appli
29	43	100.0	443	13	US-10-006-773-2	Sequence 2, Appli
30	43	100.0	443	13	US-10-006-771A-2	Sequence 2, Appli
31	43	100.0	669	17	US-10-900-928-3	Sequence 3, Appli
32	39	90.7	105	16	US-10-437-963-156937	Sequence 156937,
33	39	90.7	110	14	US-10-096-246-13	Sequence 13, Appli
34	39	90.7	239	14	US-10-096-246-32	Sequence 32, Appli
35	39	90.7	239	14	US-10-096-246-33	Sequence 33, Appli
36	39	90.7	240	14	US-10-096-246-34	Sequence 34, Appli
37	39	90.7	240	14	US-10-096-246-35	Sequence 35, Appli
38	39	90.7	240	14	US-10-096-246-36	Sequence 36, Appli
39	39	90.7	240	14	US-10-096-246-37	Sequence 37, Appli
40	39	90.7	257	14	US-10-096-246-2	Sequence 2, Appli
41	39	90.7	296	17	US-10-784-305-2	Sequence 2, Appli
42	39	90.7	622	15	US-10-378-832A-2	Sequence 2, Appli
43	38	88.4	7	14	US-10-281-479A-29	Sequence 29, Appli
44	38	88.4	7	14	US-10-160-506-5	Sequence 5, Appli
45	38	88.4	7	14	US-10-275-180A-29	Sequence 29, Appli

ALIGNMENTS

RESULT 1  
US-09-253-794-21  
; Sequence 21, Application US/09253794  
; Patent No. US0020018750A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/253,794  
; FILING DATE: 22-Feb-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/318,157  
; FILING DATE: 05-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665  
; REFERENCE/DOCKET NUMBER: 18733/464  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-253-794-21

Query Match 100.0%; Score 43; DB 9; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7  
Db 1 WTSTRHT 7

RESULT 2  
US-10-680-734-21  
; Sequence 21, Application US/10680734  
; Publication No. US20040191248A1  
; GENERAL INFORMATION:  
; APPLICANT: GOLDENBERG, DAVID M.  
; APPLICANT: HANSEN, HANS J.  
; TITLE OF INVENTION: ANTIBODY THERAPY  
; FILE REFERENCE: 40923-0051US  
; CURRENT APPLICATION NUMBER: US/10/680,734  
; CURRENT FILING DATE: 2003-10-08  
; PRIOR APPLICATION NUMBER: 60/467,161  
; PRIOR FILING DATE: 2003-05-02  
; PRIOR APPLICATION NUMBER: PCT/US02/32307  
; PRIOR FILING DATE: 2002-10-11  
; PRIOR APPLICATION NUMBER: 60/416,531  
; PRIOR FILING DATE: 2002-10-08  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn ver. 3.2  
; SEQ ID NO 21  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-10-680-734-21

Query Match 100.0%; Score 43; DB 16; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7  
Db 1 WTSTRHT 7

RESULT 3  
US-10-755-382-21  
; Sequence 21, Application US/10755382  
; Publication No. US20050089538A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, HANS J.  
; APPLICANT: ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/755,382  
; FILING DATE: 13-Jan-2004  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/253,794  
; FILING DATE: 22-Feb-1999  
; APPLICATION NUMBER: US 08/318,157  
; FILING DATE: 05-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665  
; REFERENCE/DOCKET NUMBER: 18733/464  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-10-755-382-21  
  
Query Match 100.0%; Score 43; DB 17; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 WTSTRHT 7  
Db 1 WTSTRHT 7  
  
RESULT 4  
US-09-253-794-4  
; Sequence 4, Application US/09253794  
; Patent No. US20020018750A1  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; APPLICANT: ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/253,794  
; FILING DATE: 22-Feb-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/318,157  
; FILING DATE: 05-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665



REFERENCE/DOCKET NUMBER: 18733/464  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-253-794-4

Query Match 100.0%; Score 43; DB 9; Length 106;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
Db 50 WTSTRHT 56

RESULT 5  
US-09-253-794-19  
Sequence 19, Application US/09253794  
Patent No. US20020018750A1  
GENERAL INFORMATION:  
APPLICANT: HANSEN, Hans J.  
ARMOUR, Kathryn L.

TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
MOUSE MONOCLONAL ANTIBODIES

NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/253,794  
FILING DATE: 22-Feb-1999  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/318,157  
FILING DATE: 05-OCT-1994

ATTORNEY/AGENT INFORMATION:  
NAME: SAXE, Bernhard D.  
REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 18733/464

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136

INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:

LENGTH: 106 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Query Match 100.0%; Score 43; DB 9; Length 106;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
Db 50 WTSTRHT 56

RESULT 6  
US-10-680-734-4  
Sequence 4, Application US/10680734  
Publication No. US20040191248A1  
GENERAL INFORMATION:  
APPLICANT: GOLDENBERG, DAVID M.  
APPLICANT: HANSEN, HANS J.  
TITLE OF INVENTION: ANTIBODY THERAPY  
FILE REFERENCE: 40923-0051US5  
CURRENT APPLICATION NUMBER: US/10/680,734  
CURRENT FILING DATE: 2003-10-08  
PRIOR APPLICATION NUMBER: 60/467,161  
PRIOR FILING DATE: 2003-05-02  
PRIOR APPLICATION NUMBER: PCT/US02/32307  
PRIOR FILING DATE: 2002-10-11  
PRIOR APPLICATION NUMBER: 60/416,531  
PRIOR FILING DATE: 2002-10-08  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 3.2  
SEQ ID NO 4  
LENGTH: 106  
TYPE: PRT  
ORGANISM: Mus sp.  
US-10-680-734-4

Query Match 100.0%; Score 43; DB 16; Length 106;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
Db 50 WTSTRHT 56

RESULT 7  
US-10-680-734-19  
Sequence 19, Application US/10680734  
Publication No. US20040191248A1  
GENERAL INFORMATION:  
APPLICANT: GOLDENBERG, DAVID M.  
APPLICANT: HANSEN, HANS J.  
TITLE OF INVENTION: ANTIBODY THERAPY  
FILE REFERENCE: 40923-0051US5  
CURRENT APPLICATION NUMBER: US/10/680,734  
CURRENT FILING DATE: 2003-10-08  
PRIOR APPLICATION NUMBER: 60/467,161  
PRIOR FILING DATE: 2003-05-02  
PRIOR APPLICATION NUMBER: PCT/US02/32307  
PRIOR FILING DATE: 2002-10-11  
PRIOR APPLICATION NUMBER: 60/416,531  
PRIOR FILING DATE: 2002-10-08  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 3.2  
SEQ ID NO 19  
LENGTH: 106  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-680-734-19

Query Match 100.0%; Score 43; DB 16; Length 106;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
Db 50 WTSTRHT 56

```
RESULT 8
US-10-755-382-4
; Sequence 4, Application US/10755382
; Publication No. US20050089538A1
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
;           ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
;           MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/755,382
; FILING DATE: 13-Jan-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/253,794
; FILING DATE: 22-Feb-1999
; APPLICATION NUMBER: US 08/318,157
; FILING DATE: 05-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-755-382-4
Query Match 100.0%; Score 43; DB 17; Length 106;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7
Db 50 WTSTRHT 56

RESULT 9
US-10-755-382-19
; Sequence 19, Application US/10755382
; Publication No. US20050089538A1
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
;           ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
;           MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
```

```

; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/755,382
; FILING DATE: 13-Jan-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/253,794
; FILING DATE: 22-Feb-1999
; APPLICATION NUMBER: US 08/318,157
; FILING DATE: 05-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 106 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-755-382-19
Query Match 100.0%; Score 43; DB 17; Length 106;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7
Db 50 WTSTRHT 56

RESULT 10
US-10-270-073-4
; Sequence 4, Application US/10270073
; Publication No. US20030148409A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND
; APPLICANT: CHANG, CHIEN-HSING KEN
; APPLICANT: GOLDENBERG, DAVID
; TITLE OF INVENTION: DIRECT TARGETING BINDING PROTEINS
; FILE REFERENCE: 042418/0112
; CURRENT APPLICATION NUMBER: US/10/270,073
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/328,835
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/341,881
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/345,641
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: 60/404,919
; PRIOR FILING DATE: 2002-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: hmn-14VK deduced amino acid sequence
US-10-270-073-4
Query Match 100.0%; Score 43; DB 14; Length 115;
```

```

Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRT 7
DB 50 WTSTRT 56

RESULT 11
US-10-006-771A-6
; Sequence 6, Application US/10006771A
; Publication No. US20020165360A1
; GENERAL INFORMATION:
; APPLICANT: Junghans, Richard P.
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen
; FILE REFERENCE: 002
; CURRENT APPLICATION NUMBER: US/10/006,771A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/250,090
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens and Mus sp.
US-10-006-771A-6

Query Match 100.0%; Score 43; DB 13; Length 232;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRT 7
DB 69 WTSTRT 75

RESULT 12
US-10-377-109-4
; Sequence 4, Application US/10377109
; Publication No. US20040018557A1
; GENERAL INFORMATION:
; APPLICANT: QU, ZHENGXING
; APPLICANT: HANSEN, HANS
; APPLICANT: GOLDENBERG, DAVID
; TITLE OF INVENTION: BISPECIFIC ANTIBODY POINT MUTATIONS FOR ENHANCING RATE
; FILE REFERENCE: 018733/1159
; CURRENT APPLICATION NUMBER: US/10/377,109
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/361,037
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-377-109-4

Query Match 100.0%; Score 43; DB 15; Length 232;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRT 7
DB 69 WTSTRT 75

RESULT 13
US-09-887-853-2
; Sequence 2, Application US/09887853
; Patent No. US20020168375A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; Oppermann, Hermann
; Houston, L. L.
; Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/887,853
; FILING DATE: 21-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,804
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-887-853-2

Query Match 100.0%; Score 43; DB 9; Length 250;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTSTRT 7
DB 185 WTSTRT 191

RESULT 14
US-10-683-547-2
; Sequence 2, Application US/10683547
; Publication No. US20050058638A1
; GENERAL INFORMATION:
; APPLICANT: Huston, J.
; APPLICANT: Houston, L.L.
; APPLICANT: Ring, D.
; APPLICANT: Oppermann, H.
; TITLE OF INVENTION: BIOSYNTHETIC BINDING PROTEINS FOR IMMUNO-TARGETING
; FILE REFERENCE: CIBT-P01-130
; CURRENT APPLICATION NUMBER: US/10/683,547
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US/09/558,741
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 07/831,967
; PRIOR FILING DATE: 1992-02-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 250

```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 741F8 sFv'
US-10-683-547-2
Query Match      100.0%; Score 43; DB 17; Length 250;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 WTSTRHT 7
Db      185 WTSTRHT 191
```

```
RESULT 15
US-10-270-073-6
; Sequence 6, Application US/10270073
; Publication No. US20030148409A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND
; APPLICANT: CHANG, CHIEN-HSING KEN
; APPLICANT: GOLDENBERG, DAVID
; TITLE OF INVENTION: DIRECT TARGETING BINDING PROTEINS
; FILE REFERENCE: 042418/0112
; CURRENT APPLICATION NUMBER: US/10/270,073
; PRIOR APPLICATION NUMBER: 60/328,835
; PRIOR FILING DATE: 2002-10-15
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/341,881
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/345,641
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: 60/404,919
; PRIOR FILING DATE: 2002-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: hMN-14-0 amino acid sequence
US-10-270-073-6
```

```
Query Match      100.0%; Score 43; DB 14; Length 256;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 WTSTRHT 7
Db      191 WTSTRHT 197
```

Search completed: August 9, 2005, 14:26:49  
Job time : 75.6538 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:06:08 ; Search time 15.6154 Seconds  
(without alignments)  
43.132 Million cell updates/sec

Title: US-10-089-452-28  
Perfect score: 43  
Sequence: 1 WTSTRHT 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	88.4	117	2 S42466	Ig kappa chain V r
2	38	88.4	131	2 PL0207	anti-idiotypic ant
3	38	88.4	152	2 S30751	Ig kappa chain pre
4	37	86.0	738	2 T01897	ethylene-response
5	36	83.7	461	2 B69676	alkaline phosphata
6	36	83.7	736	2 T06271	probable ethylene-
7	34	79.1	195	2 T32807	hypothetical prote
8	34	79.1	488	2 S64140	hypothetical prote
9	34	79.1	546	2 A69166	5-methylcytosine-S
10	34	79.1	745	2 T07184	subtilisin-like pr
11	34	79.1	745	2 JC6119	subtilisin-like pr
12	34	79.1	939	2 E82121	peptidase, insulin
13	32	74.4	91	2 G90732	hypothetical prote
14	32	74.4	91	2 A85583	unknown protein en
15	32	74.4	268	2 E86962	probable oxidoredu
16	32	74.4	268	2 E70551	probable oxidoredu
17	32	74.4	358	2 AF3187	flavin-dependent o
18	32	74.4	369	2 S38744	serotonin receptor
19	32	74.4	370	2 I48231	phosphopentomutase
20	32	74.4	407	2 AE1073	phosphopentomutase
21	32	74.4	407	2 F91296	phosphopentomutase
22	32	74.4	407	2 H86137	phosphopentomutase
23	32	74.4	407	2 S56507	phosphopentomutase
24	32	74.4	434	2 AG1697	hypothetical prote
25	32	74.4	572	1 DEECPC	pyruvate dehydroge
26	32	74.4	572	2 E90748	pyruvate oxidase [
27	32	74.4	572	2 A85599	pyruvate oxidase p
28	32	74.4	607	2 AB0915	probable transfera
29	32	74.4	712	2 T33231	hypothetical prote

30	32	74.4	995	2 A88483	protein C05D11.1 [
31	32	74.4	2363	2 T38841	probable pre-mRNA
32	32	74.4	2616	2 A57096	nuclei protein prec
33	31	72.1	26	2 A58955	metalloendoprotein
34	31	72.1	88	2 A30503	Ig gamma-2b chain
35	31	72.1	104	2 JC7376	hardening-inducibl
36	31	72.1	112	2 B30503	Ig gamma-2a chain
37	31	72.1	167	2 T35320	probable membrane
38	31	72.1	198	2 S30968	major tail protein
39	31	72.1	205	1 CYBOB	beta-crystallin B2
40	31	72.1	205	2 JC2009	beta-B2-crystallin
41	31	72.1	205	2 JC4964	beta-crystallin B2
42	31	72.1	207	2 A39757	beta-crystallin B2
43	31	72.1	223	1 QBE26	B2L2 protein - hu
44	31	72.1	247	2 A13555	histidine utilizat
45	31	72.1	273	2 H75393	conserved hypothet

ALIGNMENTS

RESULT 1

S42466  
Ig kappa chain V region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 21-Jan-2000  
C:Accession: S42466  
R:Shiyanov, P.A.; Respalov, I.A.; Terletskaia, H.N.; Deyev, S.M.  
submitted to the EMBL Data Library, March 1994  
A:Reference number: S42466  
A:Accession: S42466  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-117 <SHI>  
A:Cross-references: EMBL:X78108; NID:G460824; PIDN:CAA54998.1; PID:G460825  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:26-100/Domain: immunoglobulin homology <IMM>

Query Match 88.4%; Score 38; DB 2; Length 117;  
Best Local Similarity 85.7%; Pred. No. 1.3;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTSTRHT 7

Db 60 WASTRHT 66

RESULT 2

PL0207  
anti-idiotypic antibody E225, kappa chain V region precursor - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 21-Jan-2000  
C:Accession: PL0207  
R:Souchon, H.; Doyen, N.; Riottot, M.M.; Rougeon, F.; Poljak, R.J.  
Mol. Immunol. 27, 429-433, 1990  
A:Title: Nucleotide sequence of the VH, VL regions of an anti-idiotopic antibody reacting  
A:Reference number: PL0207; MUID:90309764; PMID:1973259  
A:Accession: PL0207

A:Molecule type: mRNA  
A:Residues: 1-131 <SOU>  
A>Note: This sequence corresponds to subgroup V mouse immunoglobulin light chain  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:36-110/Domain: immunoglobulin homology <IMM>  
F:44-54/Region: complementarity-determining 1  
F:70-76/Region: complementarity-determining 2  
F:109-117/Region: complementarity-determining 3  
F:116-127/Region: J region

Query Match 88.4%; Score 38; DB 2; Length 131;  
Best Local Similarity 85.7%; Pred. No. 1.5;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY      1 WTSTRHT 7
      | | | | |
Db      70 WASTRHT 76

RESULT 3
S30751
Iq kappa chain precursor V region - mouse
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Jan-2000
C:Accession: S30751
R:Grant, F.J.; Levin, S.D.; Gilbert, T.; Kindsvogel, W.
Nucleic Acids Res. 15, 5496, 1987
A>Title: Improved RNA sequencing method to determine immunoglobulin mRNA sequence.
A:Reference number: S30751; MUID:87260030; PMID:3601683
A:Accession: S30751
A:Molecule type: mRNA
A:Residues: 1-152 <GRA>
A:Cross-references: EMBL:X05877; NID:G52195; PIDN:CAA29301.1; PID:G52196
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:45-119/Domain: immunoglobulin homology <IMM>
F:138-152/Domain: C region (C-kappa) (fragment) #status predicted <CRE>

Query Match      88.4%; Score 38; DB 2; Length 152;
Best Local Similarity 85.7%; Pred. No. 1.7;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 WTSTRHT 7
      | | | | |
Db      79 WASTRHT 85

RESULT 4
T01897
ethylene-response protein ETR1 - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01897
R:Knoester, M.; Hennig, J.; Van Loon, L.C.; Bol, J.F.; Linthorst, H.J.M.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z14452
A:Accession: T01897
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-738 <KNO>
A:Cross-references: UNIPROT:O48929; EMBL:AF022727; NID:G2655060; PID:G2655061
A:Experimental source: cultivar Samsun NN
C:Superfamily: hybrid-type ethylene sensor histidine kinase; response regulator homology
C:Keywords: phosphoprotein
F:614-726/Domain: response regulator homology <RRH>
F:661/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match      86.0%; Score 37; DB 2; Length 738;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 WTSTRHT 7
      | | | | |
Db      73 WTSTAHT 79

RESULT 5
B69676
alkaline phosphatase (EC 3.1.3.1) IV precursor - Bacillus subtilis
N:Alternate names: alkaline phosphatase A (phoA); alkaline phosphomonoesterase; glycerol
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Aug-2004
C:Accession: B69676; A53375; A39096; J00808
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

```

```

Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
A:ech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauelel,
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrata, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: B69676
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-461 <KUN>
A:Cross-references: UNIPROT:P19406; GB:Z99109; GB:AL009126; NID:G2633260; PIDN:CAB12780.1
A:Experimental source: strain 168
R:Hulet, F.M.; Lee, J.; Shi, L.; Sun, G.; Chesnut, R.; Sharkova, E.; Duggan, M.F.; Kapp,
J. Bacteriol. 176, 1348-1358, 1994
A>Title: Sequential action of two-component genetic switches regulates the PHO regulon in
A:Reference number: A53375; MUID:94158639; PMID:8113174
A:Accession: A53375
A:Molecule type: DNA
A:Residues: 4-84 <HUL1>
A:Cross-references: GB:U02550; NID:G470382
R:Hulet, F.M.; Kim, E.E.; Bookstein, C.; Kapp, N.V.; Edwards, C.W.; Wyckoff, H.W.
J. Biol. Chem. 266, 1077-1084, 1991
A>Title: Bacillus subtilis alkaline phosphatases III and IV. Cloning, sequencing, and con
A:Reference number: A39096; MUID:91093215; PMID:1898729
A:Accession: A39096
A:Molecule type: DNA
A:Residues: 39-461 <HUL2>
A:Cross-references: GB:M37165; NID:G470382
R:Kapp, N.V.; Edwards, C.W.; Chesnut, R.S.; Hulet, F.M.
Gene 96, 95-100, 1990
A>Title: The Bacillus subtilis phoAIV gene: effects of in vitro inactivation on total al
A:Reference number: J00808; MUID:91092508; PMID:2125017
A:Accession: J00808
A:Molecule type: DNA
A:Residues: 39-142 <KAP>
C:Genetics:
A:Gene: phoA; phoAIV
A:Map position: 73 min
C:Superfamily: Alkaline phosphatase
C:Keywords: phosphoprotein; phosphoric monoester hydrolase
F:1-41/Domain: signal sequence #status predicted <SIG>
F:42-461/Product: alkaline phosphatase IV #status experimental <MAT>

Query Match      83.7%; Score 36; DB 2; Length 461;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 WTSTRHT 7
      | | | | |
Db      418 WTSTDHT 424

RESULT 6
T06271
probable ethylene-response protein ETR2 - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T06271; T07847
R:Lashbrook, C.C.; Tieman, D.M.; Klee, H.J.
submitted to the EMBL Data Library, January 1998
A:Description: Differential regulation of the tomato ETR gene family throughout plant dev
A:Reference number: Z15580
A:Accession: T06271
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

```

A;Title: mRNA capping enzyme. Isolation and characterization of the gene encoding mRNA 9  
A;Reference number: S31266; MUID:92250598; PMID:1315757  
A;Accession: S31267  
A;Molecule type: DNA  
A;Residues: 39-488 <SHI>  
A;Cross-references: EMBL:D10263; NID:q529646; PIDN:BAA01104.1; PID:dl001573; PID:g218413  
R;Rodriguez-Belmonte, E.; Rodriguez-Torres, A.M.; Tizon, B.; Cadahia, J.L.; Gonzalez-Siesc  
Yeast 12, 145-148, 1996  
A;Title: Sequence analysis of a 10 kb DNA fragment from yeast chromosome VII reveals a nc  
A;Reference number: S63783; MUID:96287651; PMID:8686378  
A;Accession: S63784  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-188; 'VG', '191', 'EP', '194-488 <ROD>  
A;Cross-references: EMBL:X87252  
R;Cerdan, E.  
submitted to the EMBL Data Library, May 1995  
A;Reference number: S72397  
A;Accession: S72397  
A;Molecule type: DNA  
A;Residues: 1-488 <CSW>  
A;Cross-references: EMBL:X87252; NID:g1246906; PIDN:CAA60702.1; PID:ei63176; PID:g1246906  
C;Genetics:  
A;Gene: SGD:RSM23  
A;Cross-references: SGD:S0003097  
A;Map position: 7L  
A;Note: YGL129c

Query Match 79.1%; Score 34; DB 2; Length 488;  
Best Local Similarity 71.4%; Pred. No. 39;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTSTRHT 7  
|||  
DB 98 WTNTHHT 104

RESULT 9  
A69166  
5-methylcytosine-specific restriction enzyme McrB related protein - Methanobacterium ther;  
C;Species: Methanobacterium thermoautotrophicum  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: A69166  
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F  
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.  
; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A;Reference number: A69000; MUID:98037514; PMID:9371463  
A;Accession: A69166  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-546 <MTH>  
A;Cross-references: UNIPROT:Q26601; GB:AE000833; GB:AE000666; NID:g2621559; PIDN:AAB8500  
A;Experimental source: strain Delta H  
C;Genetics:  
A;Gene: MTH501  
A;Start codon: GTG

Query Match 79.1%; Score 34; DB 2; Length 546;  
Best Local Similarity 83.3%; Pred. No. 44;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTSTRH 6  
|||  
DB 277 WTETRH 282

RESULT 10  
T07184  
subtilisin-like proteinase (EC 3.4.21.-) precursor P69B, pathogenesis-related - tomato  
C;Species: Lycopersicon esculentum (tomato)  
C;Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004

C;Accession: T07184  
R;Tornerio, P.; Conejero, V.; Vera, P.  
J. Biol. Chem. 272, 14412-14419, 1997  
A;Title: Identification of a new pathogen-induced member of the subtilisin-like processing protease family  
A;Reference number: Z15980; MUID:97306355; PMID:9162080  
A;Accession: T07184  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-745 <TOR>  
A;Cross-references: UNIPROT:Q04678; EMBL:Y10149; NID:g2230958; PIDN:CAA71234.1; PID:g2230958  
A;Experimental source: cultivar Rutgers; leaf  
C;Genetics:  
A;Gene: P69B  
C;Superfamily: subtilisin-like proteinase ag12; subtilisin homology  
C;Keywords: hydrolase; serine proteinase  
F;1-115/Domain: signal sequence and amino-terminal propeptide #status predicted <PRO>  
F;115-745/Product: subtilisin-like proteinase #status predicted <MR>  
  
Query Match 79.1%; Score 34; DB 2; Length 745;  
Best Local Similarity 71.4%; Pred. No. 62;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 WTSTRHT 7  
||| |||  
Db 729 WTSNRHS 735  
  
RESULT 11  
JC6119  
subtilisin-like proteinase (EC 3.4.21.-) - tomato  
N;Alternate names: calcium-activated endopeptidase; proteinase p69  
C;Species: Lycopersicon esculentum (tomato)  
C;Date: 23-Mar-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004  
C;Accession: JC6119  
R;Tornerio, P.; Conejero, V.; Vera, P.  
Proc. Natl. Acad. Sci. U.S.A. 93, 6332-6337, 1996  
A;Title: Primary structure and expression of a pathogen-induced protease (PR-P69) in tomato  
A;Reference number: JC6119; MUID:96270538; PMID:8692815  
A;Accession: JC6119  
A;Molecule type: mRNA  
A;Residues: 1-745 <TOR>  
A;Cross-references: UNIPROT:Q96478; EMBL:X95270; NID:g1524114; PIDN:CAA64566.1; PID:g1524114  
C;Comment: This enzyme, a secreted calcium-activated endopeptidase, is a plant pathogenesis factor  
fense against attacking pathogens.  
C;Genetics:  
A;Gene: p69  
C;Superfamily: subtilisin-like proteinase ag12; subtilisin homology  
C;Keywords: hydrolase; serine proteinase  
F;137-546/Domain: subtilisin homology #status atypical <SBT>  
  
Query Match 79.1%; Score 34; DB 2; Length 745;  
Best Local Similarity 71.4%; Pred. No. 62;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 WTSTRHT 7  
||| |||  
Db 728 WTSNRHS 734  
  
RESULT 12  
E82121  
peptidase, insulinase family VC2072 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C;Species: Vibrio cholerae  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: E82121  
R;Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Haft, D.;  
Harrison, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, J.;  
Chang, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A;Reference number: A82035; MUID:20406833; PMID:10952301  
A;Accession: E82121  
A;Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-939 <HEI>  
A;Cross-references: UNIPROT:Q9KQC8; GB:AE004281; GB:AE003852; NID:g9656616; PIDN:AAF9521E  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VC2072  
A;Map position: 1  
C;Superfamily: insulin-degrading enzyme (IDE)  
  
Query Match 79.1%; Score 34; DB 2; Length 939;  
Best Local Similarity 71.4%; Pred. No. 80;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 WTSTRHT 7  
||| |||  
Db 104 WTGTEHT 110  
  
RESULT 13  
G90732  
hypothetical protein ECs0831 [imported] - Escherichia coli (strain O157:H7, substrain RIM)  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C;Accession: G90732  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: G90732  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-91 <HAY>  
A;Cross-references: UNIPROT:Q8X863; GB:BA000007; PIDN:BA834254.1; PID:g13360290; GSPDB:G  
A;Experimental source: strain O157:H7, substrain RIMD 0503952  
C;Genetics:  
A;Gene: ECs0831  
  
Query Match 74.4%; Score 32; DB 2; Length 91;  
Best Local Similarity 83.3%; Pred. No. 16;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 WTSTRH 6  
||| |||  
Db 74 WTVTRH 79  
  
RESULT 14  
A85583  
unknown protein encoded by prophage CP-933K [imported] - Escherichia coli (strain O157:H7)  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: A85583  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,  
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: A85583  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-91 <STO>  
A;Cross-references: UNIPROT:Q8X863; GB:AE005174; NID:g12513737; PIDN:AAG55125.1; GSPDB:G  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: Z0969  
  
Query Match 74.4%; Score 32; DB 2; Length 91;  
Best Local Similarity 83.3%; Pred. No. 16;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 WTSTRH 6  
||| |||



Wed Aug 10 18:46:49 2005

DB 74 WTVTRH 79

RESULT 15

E86962  
 Probable oxidoreductase ML0429 [imported] - Mycobacterium leprae  
 C;Species: Mycobacterium leprae  
 C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
 C;Accession: E86962  
 R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H  
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
 eam, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001  
 A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
 A;Title: Massive gene decay in the leprosy bacillus.  
 A;Reference number: A86909; PMID:21128732; PMID:11234002  
 A;Accession: E86962  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-268 <STO>  
 A;Cross-references: UNIPROT:O07709; GB:AL450380; NID:G13092684; PIDN:CAC29937.1; GSPDB:G  
 C;Genetics:  
 A;Gene: ML0429  
 C;Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 74.4%; Score 32; DB 2; Length 268;  
 Best Local Similarity 71.4%; Pred. No. 51;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WTSTRHT 7  
 DB 212 WISTEHT 218

Search completed: August 9, 2005, 14:20:40  
 Job time : 17.6154 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:05:28 ; Search time 75.1154 Seconds  
(without alignments)  
47.721 Million cell updates/sec

Title: US-10-089-452-28  
Perfect score: 43  
Sequence: 1 WTSTRHT 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	90.7	743	2 Q9LWA4	Q9Lwa4 lycopersico
2	35	90.7	746	2 Q9LWA3	Q9Lwa3 lycopersico
3	39	90.7	865	2 Q6TXF5	Q6txf5 rattus norv
4	37	86.0	354	2 Q8W2M6	Q8w2m6 petunia hyb
5	37	86.0	738	1 ETR1_TOBAC	Q48929 nicotiana t
6	37	86.0	738	2 Q8W2M8	Q8w2m8 petunia hyb
7	37	86.0	738	2 Q8W2M9	Q8w2m9 petunia hyb
8	36	83.7	443	2 Q8CUS8	Q8cus8 oceanobacil
9	36	83.7	461	1 PPB4_BACSU	P19406 bacillus su
10	36	83.7	736	1 ETR2_LYCES	Q49187 lycopersico
11	35	81.4	118	2 Q77171	Q77171 plasmodium
12	35	81.4	118	2 Q8CA09	Q8ca09 mus musculu
13	35	81.4	120	2 Q8T429	Q8t429 plasmodium
14	35	81.4	121	2 Q8T533	Q8t533 plasmodium
15	35	81.4	121	2 Q868X3	Q868x3 plasmodium
16	35	81.4	121	2 Q868X5	Q868x5 plasmodium
17	35	81.4	121	2 Q9X276	Q9x276 plasmodium
18	35	81.4	122	2 Q77193	Q77193 plasmodium
19	35	81.4	122	2 Q8T4X5	Q8t4x5 plasmodium
20	35	81.4	122	2 Q9X275	Q9x275 plasmodium
21	35	81.4	123	2 Q8T332	Q8t332 plasmodium
22	35	81.4	124	2 Q77190	Q77190 plasmodium
23	35	81.4	124	2 Q6E2M4	Q6e2m4 uncultured
24	35	81.4	125	2 Q8T506	Q8t506 plasmodium
25	35	81.4	125	2 Q868Y7	Q868y7 plasmodium
26	35	81.4	125	2 Q868R8	Q868r8 plasmodium
27	35	81.4	126	2 Q868W0	Q868w0 plasmodium
28	35	81.4	126	2 Q868R6	Q868r6 plasmodium
29	35	81.4	127	2 Q8T345	Q8t345 plasmodium
30	35	81.4	127	2 Q8T420	Q8t420 plasmodium
31	35	81.4	128	2 Q9XZA5	Q9xza5 plasmodium

32	35	81.4	129	2 Q95Z31	Q95z31 plasmodium
33	35	81.4	129	2 Q9NHL3	Q9nhl3 plasmodium
34	35	81.4	130	2 Q77184	Q77184 plasmodium
35	35	81.4	130	2 Q8T4V6	Q8t4v6 plasmodium
36	35	81.4	130	2 Q9XZ84	Q9xz84 plasmodium
37	35	81.4	131	2 Q8STF4	Q8stf4 plasmodium
38	35	81.4	131	2 Q8T4U7	Q8t4u7 plasmodium
39	35	81.4	131	2 Q8T4X9	Q8t4x9 plasmodium
40	35	81.4	131	2 Q8T577	Q8t577 plasmodium
41	35	81.4	131	2 Q868Z2	Q868z2 plasmodium
42	35	81.4	132	2 Q8T355	Q8t355 plasmodium
43	35	81.4	133	2 Q964S4	Q964s4 plasmodium
44	35	81.4	133	2 Q868R3	Q868r3 plasmodium
45	35	81.4	133	2 Q868R4	Q868r4 plasmodium

ALIGNMENTS

RESULT 1  
Q9LWA4 PRELIMINARY; PRT; 743 AA.  
AC Q9LWA4;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Subtilisin-like protease.  
GN Name=P69E;  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; asterids;  
OC lamids; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=leaf;  
RX MEDLINE=20989714; PubMed=10631250; DOI=10.1104/pp.122.1.67;  
RA Jordá L., Conejero V., Vera P.;  
RT "Characterization of two differentially regulated genes (P69E and P69F) encoding new members of the subtilisin-like protease clan from tomato plants";  
RL Plant Physiol. 122:67-76(2000).  
DR EMBL; Y18931; CAB67119.1; -.  
DR HSSP; Q99405; IMPT.  
DR MEROPS; S08.006; -.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004289; F:subtilase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR003137; PA.  
DR InterPro; IPR000209; Pept\_S8\_S53.  
DR InterPro; IPR009020; Prot\_inh\_propept.  
DR InterPro; IPR010259; Prot\_inh\_S8A.  
DR Pfam; PF02225; PA; 1.  
DR Pfam; PF00082; Peptidase\_S8; 1.  
DR Pfam; PF05922; Subtilisin\_N; 1.  
DR PRINTS; PR00723; SUBTILISIN.  
DR PROSITE; PS00136; SUBTILASE\_ASP; UNKNOWN\_1.  
DR PROSITE; PS00138; SUBTILASE\_SER; UNKNOWN\_1.  
KW Protease.  
SQ SEQUENCE 743 AA; 79133 MW; 63BE2F92B2510BCA CRC64;

Query Match 90.7%; Score 39; DB 2; Length 743;  
Best Local Similarity 85.7%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTSTRHT 7  
Db 727 WTSTRHS 733

RESULT 2  
Q9LWA3 PRELIMINARY; PRT; 746 AA.  
ID Q9LWA3

AC Q91WA3;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Subtilisin-like protease.  
 GN Name=P69F;  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Leaf;  
 RC MEDLINE=20098714; PubMed=10631250; DOI=10.1104/pp.122.1.67;  
 RA Jorda L., Conejero V., Vera P.;  
 RT "Characterization of two differentially regulated genes (P69E and P69F) encoding new members of the subtilisin-like protease clan from tomato plants.";  
 RL Plant Physiol. 122:67-76(2000).  
 DR EMBL; Y18932; CAB67120.1; -;  
 DR HSSP; P29599; 1ST3.  
 DR MEROPS; S08.006; -;  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004289; F:subtilase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001317; PA.  
 DR InterPro; IPR002029; Pept\_S8\_S53.  
 DR InterPro; IPR009020; Prot\_inh\_propept.  
 DR InterPro; IPR010259; Prot\_inh\_S8A.  
 DR Pfam; PF02225; PA; 1.  
 DR Pfam; PF00082; Peptidase\_S8; 1.  
 DR Pfam; PF05922; Subtilisin\_N; 1.  
 DR PRINTS; PR00723; SUBTILISIN.  
 DR PROSITE; PS00136; SUBTILASE\_ASP; UNKNOWN\_1.  
 DR PROSITE; PS00138; SUBTILASE\_SER; UNKNOWN\_1.  
 KW Protease.  
 SQ SEQUENCE 746 AA; 79062 MW; F2B9D6B7EA0EA4AB CRC64;  
 Query Match 90.7%; Score 39; DB 2; Length 746;  
 Best Local Similarity 85.7%; Pred. No. 55;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 WTSTRHT 7  
 Db 730 WTSTRHS 736  
 RESULT 3  
 O6TXF5 PRELIMINARY; PRT; 865 AA.  
 AC Q6TXF5;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE LRRGT00044.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Sprague-Dawley;  
 RA Xu C.S., Chang C.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,  
 RA Yang K.J., Zhao L.F., Ma H., Wang L., Wang S.F., Xing X.K., Shen G.M.,  
 RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY383699; AAQ96257.1; -;  
 DR InterPro; IPR001611; LRR.  
 DR Pfam; PF00560; LRR\_1; 2.  
 DR PRINTS; PR00019; LEURICHRPT.  
 SQ SEQUENCE 865 AA; 95988 MW; 43BE96C199460485 CRC64;

Query Match 90.7%; Score 39; DB 2; Length 865;  
 Best Local Similarity 85.7%; Pred. No. 63;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 WTSTRHT 7  
 Db 43 WSSTRHT 49  
 RESULT 4  
 Q8W2MG PRELIMINARY; PRT; 354 AA.  
 ID Q8W2MG  
 AC Q8W2MG;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Ethylene receptor.  
 GN Name=ETR3;  
 OS Petunia hybrida (Petunia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Petunia.  
 OX NCBI\_TaxID=4102;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP Lai J.P., Shaw J.F.;  
 RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF145975; AAL40904.1; -;  
 DR HSSP; P43333; 1DCF.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0000156; F:two-component response regulator activity; IEA.  
 DR GO; GO:0007600; P:sensory perception; IEA.  
 DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.  
 DR Pfam; PF01590; GAF; 1.  
 DR Pfam; PF000072; Response\_reg; 1.  
 DR ProDom; PD000039; Response\_reg; 1.  
 DR SMART; SM00448; REC; 1.  
 DR PROSITE; PS0110; RESPONSE\_REGULATORY; 1.  
 KW Phosphorylation; Receptor; Sensory transduction.  
 SQ SEQUENCE 354 AA; 39927 MW; A563EB1F10D2D1ED CRC64;  
 Query Match 86.0%; Score 37; DB 2; Length 354;  
 Best Local Similarity 85.7%; Pred. No. 61;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 WTSTRHT 7  
 Db 73 WTSTAHT 79  
 RESULT 5  
 ETR1\_TOBAC STANDARD; PRT; 738 AA.  
 ID ETR1\_TOBAC  
 AC O48929;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Ethylene receptor (EC 2.7.3.-) (NT-ETR1).  
 GN Name=ETR1;  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=cv. Samsun NN; TISSUE=Leaf;  
 RC Knoester M., Hennig J., van Loon L.C., Bol J.F., Linthorst H.J.M.;  
 RA "Isolation and characterization of a tobacco cDNA encoding an ETR1 homolog.";  
 RT (et) Plant Gene Register PGR97-188.  
 CC -!- FUNCTION: May act early in the ethylene signal transduction

pathway, possibly as an ethylene receptor, or as a regulator of the pathway (By similarity).

-!- COFACTOR: Binds 1 copper ion per dimer (By similarity).

-!- SUBUNIT: Homodimer; disulfide-linked (By similarity).

-!- SUBCELLULAR LOCATION: Endoplasmic reticulum (By similarity).

-!- INDUCTION: Constitutive expression. Not induced by senescence, wounding, application of ethephon or infection with virus.

-!- PTM: Activation probably requires a transfer of a phosphate group between a His in the transmitter domain and an Asp of the receiver domain (By similarity).

-!- SIMILARITY: Belongs to the ethylene receptor family.

-!- SIMILARITY: Contains 1 histidine kinase domain.

-!- SIMILARITY: Contains 1 response regulatory domain.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; AF022727; AAB97160.1; --

PIR; T01897; T01897.

HSSP; P49333; 1DCF.

InterPro; IPR003594; ATPbind ATPase.

InterPro; IPR004358; Bact\_sens\_pr\_C.

InterPro; IPR011006; CheY\_like.

InterPro; IPR003018; GAF.

InterPro; IPR003661; His\_kinA\_N.

InterPro; IPR005467; His\_kinase.

InterPro; IPR001789; Response\_reg.

Pfam; PF01590; GAF; 1.

Pfam; PF02518; HATPase\_c; 1.

Pfam; PF00512; HsKA; 1.

Pfam; PF00072; Response\_reg; 1.

PRINTS; PR00344; BCTRLSENSOR.

ProDom; PD000039; Response\_reg; 1.

SMART; SM00065; GAF; 1.

SMART; SM00387; HATPase\_c; 1.

SMART; SM00388; HsKA; 1.

SMART; SM00448; REC; 1.

PROSITE; PS0109; HIS\_KIN; 1.

PROSITE; PS0110; RESPONSE\_REGULATORY; 1.

Copper; Kinase; Metal-binding; Phosphorylation; Receptor; Sensory transduction; Transferase; Transmembrane; Two-component regulatory system.

FT DOMAIN 348 585 Histidine kinase.

FT DOMAIN 613 730 Response regulatory.

FT TRANSMEM 22 42 Potential.

FT TRANSMEM 53 73 Potential.

FT TRANSMEM 91 111 Potential.

FT METAL 64 64 Copper (By similarity).

FT METAL 68 68 Copper (By similarity).

FT DISULFID 3 3 Interchain (By similarity).

FT DISULFID 5 5 Interchain (By similarity).

FT MOD\_RES 351 351 Phosphohistidine (by autocatalysis) (By similarity).

FT MOD\_RES 661 661 4-azapartylphosphate (By similarity).

SQ SEQUENCE 738 AA; 82294 MW; D6CSF7B16F16A7DE CRC64;

Query Match 86.0%; Score 37; DB 1; Length 738;

Best Local Similarity 85.7%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WTSTRHT 7

Db 73 WTSTAHT 79

RESULT 6

Q8W2M8 PRELIMINARY; PRT; 738 AA.

ID Q8W2M8

Q8W2M8 PRELIMINARY; PRT; 738 AA.

AC Q8W2M8

Q8W2M8; 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Ethylene receptor.

GN Name=ETR1-2;

OS Petunia hybrida (Petunia).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;

OC Lamiales; Solanales; Solanaceae; Petunia.

OX NCBI\_TaxID=4102;

RN [1]

RP SEQUENCE FROM N.A.

RA Lai J.P., Shaw J.F.;

RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: Contains 1 histidine kinase domain.

DR EMBL; AF145973; AAL40902.1; --

DR HSSP; P49333; 1DCF.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0016301; F:kinase activity; IEA.

DR GO; GO:0004872; F:receptor activity; IEA.

DR GO; GO:0000156; F:two-component response regulator activity; IEA.

DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.

DR GO; GO:0007600; P:sensory perception; IEA.

DR GO; GO:0001600; P:two-component signal transduction system (p. . .; IEA.

DR Pfam; PF01590; GAF; 1.

DR Pfam; PF02518; HATPase\_c; 1.

DR Pfam; PF00512; HsKA; 1.

DR Pfam; PF00072; Response\_reg; 1.

DR PRINTS; PR00344; BCTRLSENSOR.

DR ProDom; PD000039; Response\_reg; 1.

DR SMART; SM00065; GAF; 1.

DR SMART; SM00387; HATPase\_c; 1.

DR SMART; SM00388; HsKA; 1.

DR SMART; SM00448; REC; 1.

DR PROSITE; PS0109; HIS\_KIN; 1.

DR PROSITE; PS0110; RESPONSE\_REGULATORY; 1.

KW Phosphorylation; Receptor; Sensory transduction.

SQ SEQUENCE 738 AA; 82267 MW; 39617CEA365P821 CRC64;

Query Match 86.0%; Score 37; DB 2; Length 738;

Best Local Similarity 85.7%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WTSTRHT 7

Db 73 WTSTAHT 79

RESULT 7

Q8W2M9 PRELIMINARY; PRT; 738 AA.

ID Q8W2M9

Q8W2M9 PRELIMINARY; PRT; 738 AA.

AC Q8W2M9

Q8W2M9; 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Ethylene receptor.

GN Name=ETR1-1;

OS Petunia hybrida (Petunia).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;

OC Lamiales; Solanales; Solanaceae; Petunia.

OX NCBI\_TaxID=4102;

RN [1]

RP SEQUENCE FROM N.A.

RA Lai J.P., Shaw J.F.;

RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: Contains 1 histidine kinase domain.

DR EMBL; AF145972; AAL40901.1; --

DR HSSP; P49333; 1DCF.

DR GO; GO:0016020; C:membrane; IEA.

```

DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0036301; F:kinase activity; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0000156; F:two-component response regulator activity; IEA.
DR GO: GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO: GO:0007600; P:sensory perception; IEA.
DR GO: GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR Pfam; PF01590; GAF; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HisKA; 1.
DR Pfam; PF00072; Response_reg; 1.
DR PRINTS; PR00344; BCTRLSNSOR.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00065; GAF; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HisKA; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS01109; HIS_KIN; 1.
DR PROSITE; PS01110; RESPONSE_REGULATORY; 1.
KW Phosphorylation; Receptor; Sensory transduction.
SQ SEQUENCE 738 AA; 82407 MW; 96E12975A89ED710 CRC64;

Query Match      86.0%; Score 37; DB 2; Length 738;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTSTRTHT 7
Db 73 WTSTAHT 79

RESULT 8
Q8CUS8 PRELIMINARY; PRT; 443 AA.
AC Q8CUS8;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Alkaline phosphatase (EC 3.1.3.1).
OS OrderedLocusNames=OB1023;
GN Oceanobacillus iheyensis;
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]_TaxID=182710;
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831;
RX MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkf526;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935 (2002).
CC -!- SIMILARITY: Belongs to the alkaline phosphatase family.
DR EMBL; AP004596; BAC12985.1; -.
DR HSP; P00634; 1A7A.
DR GO: GO:0004035; F:alkaline phosphatase activity; IEA.
DR GO: GO:0008152; F:metabolism; IEA.
DR InterPro; IPR001952; Alk phosphatse.
DR Pfam; PF00245; Alk phosphatase; 1.
DR PRINTS; PR00113; ALKPHPTASE.
DR SMART; SM00098; alkPPC; 1.
DR PROSITE; PS00123; ALKALINE_PHOSPHATASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 443 AA; 48148 MW; B1C5B600116F4500 CRC64;

Query Match      83.7%; Score 36; DB 2; Length 443;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTSTRTHT 7
Db 401 WTSTGHT 407

```

```

RESULT 9
PPB4_BACSU STANDARD; PRT; 461 AA.
AC PPB4_BACSU
ID P19406;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alkaline phosphatase IV precursor (EC 3.1.3.1) (Aphase IV).
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=94156839; PubMed=8111174;
RA Hulet F.M., Lee J., Shi L., Sun G., Chesnut R., Sharkova E.,
RA Duggan M.F., Kapp N.;
RA "Sequential action of two-component genetic switches regulates the PHO
RT regulon in Bacillus subtilis.";
RL J. Bacteriol. 176:1348-1358(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98240224; PubMed=9579061;
RA Noback M.A., Holsappel S., Kiewiet R., Terpstra P., Wambutt R.,
RA Wedler H., Venema G., Bron S.;
RT "The 172 kb prkA-addAB region from 83 degrees to 97 degrees of the
RT Bacillus subtilis chromosome contains several dysfunctional genes, the
RT glvB marker, many genes encoding transporter proteins, and the
RT ubiquitous hit gene.";
RL Microbiology 144:859-875(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerger P.T.,
RA Etian K.-D., Errington J., Rabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Katamata D., Kasahara Y., Klaerr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kunano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Roy M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadale Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassartotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzenegger T., Winters P., Wipat A., Yamamoto K., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256(1997).
RN [4]
RP SEQUENCE OF 399-461 FROM N.A.
RC STRAIN=168 / JH642;

```

```

RX MEDLINE=91093215; PubMed=1898729;
RA Hulett F.M., Kim M.E., Bookstein C., Kapp N.V., Edwards C.W.,
RA Wyckoff H.W.;
RT "Bacillus subtilis alkaline phosphatases III and IV. Cloning,
RT sequencing, and comparisons of deduced amino acid sequence with
RT Escherichia coli alkaline phosphatase three-dimensional structure."
RL J. Biol. Chem. 266:1077-1084 (1991).
RN [5]
RP SEQUENCE OF 39-142 FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=91092508; PubMed=2125017; DOI=10.1016/0378-1119(90)90346-S;
RA Kapp N.V., Edwards C.W., Chesnut R.S., Hulett F.M.;
RT "The Bacillus subtilis phoAIV gene: effects of in vitro inactivation
RT on total alkaline phosphatase production."
RL Gene 96:95-100 (1990).
RN [6]
RP SEQUENCE OF 42-63.
RX MEDLINE=90130309; PubMed=2105301;
RA Hulett F.M., Bookstein C., Jensen K.;
RT "Evidence for two structural genes for alkaline phosphatase in
RT Bacillus subtilis."
RL J. Bacteriol. 172:735-740 (1990).
RN [7]
RP SEQUENCE OF 433-461 FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=94193548; PubMed=8144469;
RA Beall B.W., Moran C.P. Jr.;
RT "Cloning and characterization of spoVR, a gene from Bacillus subtilis
RT involved in spore cortex formation."
RL J. Bacteriol. 176:2003-2012 (1994).
CC -1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC alcohol + phosphate.
CC -1- COFACTOR: Metalloenzyme containing two zinc atoms and a magnesium
CC ion.
CC -1- SUBUNIT: Monomer.
CC -1- SIMILARITY: Belongs to the alkaline phosphatase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U02550; AAA18323.1; -
CC EMBL; Y14082; CAA74486.1; -
CC EMBL; Z99109; CAB12780.1; -
CC EMBL; L26337; AAA22812.1; -
CC PIR; B69676; B69676.
CC HSP; P00634; IAJA.
CC Subtilisin; BG10183; phoA.
CC InterPro; IPR001952; Alk_phosphatase.
CC Pfam; PF00245; Alk_phosphatase; 1.
CC PRINTS; PR00113; ALKPHPTASE.
CC SMART; SM00098; alkPPC; 1.
CC PROSITE; PS00123; ALKALINE PHOSPHATASE; 1.
CC Complete proteome; Direct protein sequencing; Hydrolase; Magnesium;
CC Multigene family; Phosphorylation; Signal; Zinc.
CC SIGNAL
CC CHAIN 42 461 Alkaline phosphatase IV.
CC ACT_SITE 108 108 Phosphoserine intermediate (By
CC similarity).
CC CONFLICT 50 50 R -> K (in Ref. 6).
CC SEQUENCE 461 AA; 50273 MW; A2ABD109026A2BFE CRC64;
Query Match 83.7%; Score 36; DB 1; Length 461;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 WTSTRHT 7
Db 418 WTSTDHT 424

```

```

RESULT 10
ID ETR2 LYCES STANDARD; PRT; 736 AA.
AC O49187; Q41343;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ethylene receptor 2 (EC 2.7.3.-) (LeETR2).
GN Name=ETR2;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Rutgers; TISSUE=Root;
RX MEDLINE=98388654; PubMed=9721682;
RA Lashbrook C.C., Tieman D.M., Klee H.J.;
RT "Differential regulation of the tomato ETR gene family throughout
RT plant development."
RL Plant J. 15:243-252 (1998).
RN [2]
RP SEQUENCE OF 31-736 FROM N.A.
RC STRAIN=cv. UC92B;
RA Zhou D., Mattoo A.K., Tucker M.L.;
RT "Molecular cloning of a tomato cDNA encoding an ethylene receptor."
RL (er) Plant Gene Register PGR96-015.
CC -1- FUNCTION: May act early in the ethylene signal transduction
CC pathway, possibly as an ethylene receptor, or as a regulator of
CC the pathway (By similarity).
CC -1- COFACTOR: Binds 1 copper ion per dimer (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum (By similarity).
CC -1- TISSUE SPECIFICITY: Leaves, flowers and fruits.
CC -1- DEVELOPMENTAL STAGE: Induced in imbibing seeds prior to
CC germination and down-regulated in elongating seedlings and
CC senescing leaf petioles.
CC -1- INDUCTION: Not induced by ethylene.
CC -1- PTM: Activation probably requires a transfer of a phosphate group
CC between a His in the transmitter domain and an Asp of the receiver
CC domain (By similarity).
CC -1- SIMILARITY: Belongs to the ethylene receptor family.
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF043085; AAC02214.1; -
CC EMBL; U47279; AAB39386.1; -
CC PIR; T06271; T06271.
CC HSP; P49333; LDCP.
CC InterPro; IPR003594; ATPbind_ATPase.
CC InterPro; IPR004358; Bact_sens_pr_C.
CC InterPro; IPR011006; CheY_like.
CC InterPro; IPR003018; GAF.
CC InterPro; IPR009082; His_kin_homodim.
CC InterPro; IPR003661; His_kin_N.
CC InterPro; IPR005467; His_kinase.
CC InterPro; IPR001789; Response_reg.
CC Pfam; PF01590; GAF; 1.
CC Pfam; PF02518; HATPase_c; 1.
CC Pfam; PF00512; HsKA; 1.
CC Pfam; PF00072; Response_reg; 1.
CC PRINTS; PR00344; BCTRUSENSOR.

```

```

DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00065; GAF; 1.
DR SMART; SM00387; HATPase.c; 1.
DR SMART; SM00388; Hiska; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Copper; Kinase; Metal-binding; Multigene family; Phosphorylation;
KW Receptor; Sensory transduction; Transferase; Transmembrane;
KW Two-component regulatory system..
FT DOMAIN 348 585 Histidine kinase.
FT DOMAIN 613 730 Response regulatory.
FT TRANSMEM 22 42 Potential.
FT TRANSMEM 53 73 Potential.
FT TRANSMEM 94 114 Potential.
FT METAL 64 64 Copper (By similarity).
FT METAL 68 68 Copper (By similarity).
FT DISULFID 3 3 Interchain (By similarity).
FT DISULFID 5 5 Interchain (By similarity).
FT MOD_RES 351 351 Phosphohistidine (by autocatalysis) (By
similarity)
FT MOD_RES 661 661 4-asparylphosphate (By similarity).
FT CONFLICT 474 474 K -> E (in Ref. 2).
SQ SEQUENCE 736 AA; 81701 MW; A804C61EBC190320 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 736;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTSTRT 7
Db 73 WTSTRT 79

RESULT 11
O77171 ID O77171 PRELIMINARY; PRT; 118 AA.
AC O77171;
DC 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE Erythrocyte membrane protein 1 type c (Fragment).
GN Name=var;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Tiwari B.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF084579; AAC62721.1; -.
FT NON_TER 1 1
FT NON_TER 118 118
SQ SEQUENCE 118 AA; 13893 MW; 8203D952A1F7AE5 CRC64;

Query Match 81.4%; Score 35; DB 2; Length 118;
Best Local Similarity 71.4%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTSTRT 7
Db 57 WTANRHT 63

RESULT 12
Q8CA09 ID Q8CA09 PRELIMINARY; PRT; 118 AA.
AC Q8CA09;
DC 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched
library, clone:A430041E03 product:hypothetical protein, full insert

```

```

DE sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6875(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA The PANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK039996; BAC30494.1; -.
KW Hypothetical protein.
SQ SEQUENCE 118 AA; 13454 MW; B41AF4CF0C512250 CRC64;

Query Match 81.4%; Score 35; DB 2; Length 118;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```



Qy 1 WTSTRH 6  
||:||||  
Db 4 WTATRH 9

RESULT 13

Q8T4Z9 PRELIMINARY; PRT; 120 AA.

1D	Q81429	PREMILINARIIN, FR,	120 AA.
AC	Q81429	01-JUN-2002 (TrEMBLrel. 21, Created)	
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Erythrocyte membrane protein 1 (Fragment).		
DE	Name=var;		
GN	Plasmodium falciparum.		
OS	Plasmodium falciparum.		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
ON	NCBI_TaxID=5833;		
OX	{1}		
PN			

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847

RN [2]  
RP SEQUENCE FROM N.A.  
RA Fowler E., Peters J., Gattton M., Chen N., Cheng Q.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.

	EMBL	AY054911; AAL11275.1;	-.
FT	NON_TER	1	1
FT	NON_TER	120	120
SO	SEQUENCE	120 AA:	13669 MW; 1E45687DC7CEAA62 CRC64;

Query Match 81.4%; Score 35; DB 2; Length 120;  
Best Local Similarity 71.4%; Pred. No. 49;  
Matches 5: Conservative 1: Mismatches 1: Indels

Qy 1 WTSTRHT 7  
||: |||  
pb 60 WTANRHT 6

RESULT 14

RES001 14  
Q8T533  
ID Q8T533  
PRELIMINARY;  
PRT; 121 AA.

AC	Q8T533;
AD	01-JUN-2002 (TrEMBLrel. 21, Created)
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Erythrocyte membrane protein 1 (Fragment).
GN	Name=var;
OS	Plasmodium falciparum.
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX	NCBI_TaxID=5833;
	[1]
RN	SEQUENCE FROM N.A.
RP	MEDLINE=21839615; PubMed=11849711; DOI=10.1016/S0166-6851(01)00443-1;
RX	Fowler E.V., Peters J.M., Gatton M.L., Chen N., Cheng Q.;
RA	"Genetic diversity of the DBLalpha region in Plasmodium falciparum var
RT	genes among Asia-Pacific isolates.;"
RL	Mol. Biochem. Parasitol. 120:117-126(2002).

RP SEQUENCE FROM N.A.  
 RA Fowler E., Peters J., Gatton M., Chen N., Cheng Q.;  
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY054869; AAL11233.1; -.  
 NCBI: AY054869.1; AAL11233.1; -.  
 MOL. BIOCHEM. PARASITOL. 120:117-120(2002).  
 [2]

DATA	DATE, TIME, USER, FILE, LINE	NON_TER	1	1
FT		NON_TER	121	121
FT		NON_TER	121	121
SO		SEQUENCE	121 AA;	14168 MW; AB8A9D5D2E6BE8D40 CRC64;

Query Match 81.4%; Score 35; DB 2; Length 121;  
Best Local Similarity 71.4%; Pred. NO. 49;  
Matches 5; Conservative 1; Mismatches 1; Indels

Qy 1 WTSTRHT 7  
||: |||  
pb 60 WTANRHT 66

RESULT 15

REC'D 13  
Q868X3  
ID Q868X3  
PRELIMINARY;  
PRT: 121 AA.

AC Q868Y3; (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE DYE Erythrocyte membrane protein 1 variant 1 (fragment).  
GN Names-var;  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
NCBI\_TaxID=5833;  
[1]  
DB SEQUENCE FROM NCBI

157  
AN  
RP  
RX  
RX  
RA  
RT  
RT  
RT  
RT  
PL

SEQUENCE FROM N.A.  
MEDLINE-22561043; PubMed13672527; DOI=10.1016/S0166-6851(03)00004-5;  
Winter G., Chen Q., Flick K., Krensner P., Fernandez V., Wahlgren M.;  
"The 3Dvar5.2 (var COMMON) type var gene family is commonly expressed  
in non-placental Plasmodium falciparum malaria.";  
PL Mol. Biochem. Parasitol. 127:179-191(2003).

AD	NON_TER	121	121	AA: 14337 MW: A24B9332DB2218740 CRC64:
DR	NON_TER	1	1	SEQUENCE 121 AA: 14337 MW: A24B9332DB2218740 CRC64:
FT	NON_TER	1	1	EMBL: AF528130; AA085747.1; -
FT	NON_TER	1	1	EMBL: AF528130; AA085747.1; -

Query Match 81.4%; Score 35; DB 2; Length 121;  
Best Local Similarity 71.4%; Pred. No. 49;  
Matches 5; Conservative 1; Mismatches 1; Indels

Qy 1 WTSTRHT 7  
||: |||  
pb 65 WTANRHT 71

Search completed: August 9, 2005, 14:19:36  
Job time : 76.1154 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:02:28 ; Search time 93.5385 Seconds  
(without alignments)  
33.078 Million cell updates/sec

Title: US-10-089-452-29  
Perfect score: 41  
Sequence: 1 QQYSSSPT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16dec04:.\*  
1: Geneseq1980s:.\*  
2: Geneseq1990s:.\*  
3: Geneseq2000s:.\*  
4: Geneseq2001s:.\*  
5: Geneseq2002s:.\*  
6: Geneseq2003as:.\*  
7: Geneseq2003bs:.\*  
8: Geneseq2004s:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	41	100.0	8	4	AAB86086	Aab86086 H. pylori
2	41	100.0	8	4	AAB86054	Aab86054 H. pylori
3	41	100.0	106	4	AAB86108	Aab86108 H. pylori
4	41	100.0	106	4	AAB86070	Aab86070 H. pylori
5	37	90.2	8	5	ABP82398	Abp82398 Human imm
6	37	90.2	107	5	ABP07229	Abp07229 Anti-IL-4
7	37	90.2	108	4	AAB62755	Abp62755 Human HIV
8	37	90.2	132	5	ABP62189	Abp62189 Human imm
9	37	90.2	384	4	AAM24101	Aam24101 Human EST
10	37	90.2	384	4	AAU14462	Aau14462 Human nov
11	37	90.2	384	4	AAU14463	Aau14463 Human nov
12	37	90.2	384	4	AAU14461	Aau14461 Human nov
13	37	90.2	384	4	AAU14464	Aau14464 Human nov
14	37	90.2	384	8	ADH80782	Adh80782 Human pol
15	37	90.2	384	8	ADH80779	Adh80779 Human pol
16	37	90.2	384	8	ADH80781	Adh80781 Human pol
17	37	90.2	385	8	ADH80780	Adh80780 Human pol
18	36	87.8	9	8	ADO39439	Ado39439 Novel hum
19	36	87.8	9	8	ADO39457	Ado39457 Novel hum
20	36	87.8	9	8	ADO39451	Ado39451 Novel hum
21	36	87.8	120	8	ADO39421	Ado39421 Human AB1
22	36	87.8	120	8	ADO39429	Ado39429 Human AB1
23	34	82.9	257	8	ADMI6744	Admi6744 PERS-P1-H
24	33	80.5	40	6	ABJ25312	Abj25312 SAC isola
25	33	80.5	40	6	ABJ25319	Abj25319 SAC isola

26	33	80.5	251	5	ABP44303	Abp44303 Human BLY
27	33	80.5	251	7	ADG95130	Adg95130 Single ch
28	33	80.5	283	8	ADF85426	Adf85426 V1228cFv
29	33	80.5	644	6	ABR43313	AbR43313 Canis fam
30	33	80.5	665	8	ADN41896	Adn41896 Amino aci
31	33	80.5	665	8	ADN41900	Adn41900 Amino aci
32	33	80.5	942	8	ADR24017	Adr24017 Human epi
33	33	80.5	1145	8	ABM83039	Abm83039 Human dia
34	33	80.5	1155	6	ABR54203	AbR54203 Human NOV
35	33	80.5	1188	8	ADO28625	Ado28625 Human EGF
36	33	80.5	1210	3	AAV50616	Aay50616 Human EGF
37	33	80.5	1210	3	AAB19259	Aab19259 Amino aci
38	33	80.5	1210	4	AAB68420	Aab68420 Amino aci
39	33	80.5	1210	5	AAE23019	Aae23019 Human Her
40	33	80.5	1210	5	AAM50768	Aam50768 Human epi
41	33	80.5	1210	5	ABP54803	Abp54803 Human COP
42	33	80.5	1210	5	ABP51768	Abp51768 Human epi
43	33	80.5	1210	6	ABR58530	AbR58530 Human epi
44	33	80.5	1210	6	ABR54204	AbR54204 Human NOV
45	33	80.5	1210	7	ADA37265	Ada37265 Human epi

ALIGNMENTS

RESULT 1  
AAB86086  
ID AAB86086 standard; peptide; 8 AA.  
XX  
AC AAB86086;  
XX  
DT 17-JUL-2001 (first entry)  
XX  
DE H. pylori catalase derived antibody HP25/6m/1B5 light chain CDR3.  
XX  
KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;  
KW acid-resistant microorganism; complementarity determining region; CDR;  
KW feces; heavy chain; light chain.  
XX  
OS Unidentified.  
XX  
PN WO200127612-A2.  
XX  
PD 19-APR-2001.  
XX  
PF 12-OCT-2000; 2000WO-EP010057.  
XX  
PR 12-OCT-1999; 99EP-00120351.  
PR 16-MAR-2000; 2000EP-00105592.  
PR 31-MAR-2000; 2000EP-00107028.  
PR 10-MAY-2000; 2000EP-00110110.  
PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
XX  
PI Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;  
XX  
DR N-PSDB; AAF88113.  
XX  
PT Detecting infections by acid-resistant microorganisms, particularly for  
PT diagnosing Helicobacter pylori, comprises immunochromatographic detection  
PT of antigen in feces.  
XX  
PS Claim 26; Page 26; 90pp; German.  
XX  
CC This invention describes a novel method for detecting infection by an  
CC acid-resistant microorganism (A), in a mammal, using  
CC immunochromatography. The method is used to diagnose infection by an acid  
CC -resistant microorganism (A), in a mammal, such as Helicobacter,  
CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
CC H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,  
CC inexpensive and non-invasive, and may indicate the stage of infection. A  
CC test strip used in the method may include a filter to eliminate particles

CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence represents a  
 CC complementarity determining region (CDR) from an antibody raised against  
 CC the H. pylori catalase or beta-urease antigen which is used to illustrate  
 CC the method of the invention

XX Sequence 8 AA;

Query Match 100.0%; Score 41; DB 4; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQYSSSPT 8  
 |||||  
 Db 1 QQYSSSPT 8

RESULT 2  
 AAB86054  
 ID AAB86054 standard; peptide; 8 AA.

XX AC AAB86054;

XX DT 17-JUL-2001 (first entry)

XX DE H. pylori catalase derived antibody HP25/6m/1B5 light chain CDR3.

XX KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
 infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
 KW antibacterial; complementarity determining region.

XX OS Unidentified.

XX PN WO200127613-A2.

XX PD 19-APR-2001.

XX PF 12-OCT-2000; 2000WO-EP010058.

XX PR 12-OCT-1999; 99EP-00120351.

XX PR 16-MAR-2000; 2000EP-00105592.

XX PR 31-MAR-2000; 2000EP-00107028.

XX PR 10-MAY-2000; 2000EP-00110110.

XX PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX PI Reiter C, Cullmann G, Heppner P, Ringeis A, Mueller H, Haindl E;

XX DR WPI; 2001-282087/29.

XX DR N-PSDB; AAF88056.

XX PT Detecting infections by acid-resistant microorganisms, particularly for  
 diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
 sample.

XX PS Claim 21; Page 16; 89pp; German.

XX This invention describes a novel method for detecting, in a mammal,  
 CC infection by an acid-resistant microorganism (A) which comprises reacting  
 CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
 CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
 CC complex is formed with Ag, and the formation of a complex detected. R are  
 CC specific for an Ag which, after passage through the intestines, at least  
 CC in some mammals, retains a native (or corresponding) structure against  
 CC which the mammal produces antibodies (when immunized or infected with  
 CC (A), or its extracts, lysates or derived proteins (or fragments) or  
 CC synthetic peptides). The products of the invention have antibacterial  
 CC activity. The method is used to diagnose infection by Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the  
 CC progress of treatment. Receptors, particularly antibodies, directed

CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an  
 CC infection. This sequence represents a complementarity determining region  
 CC (CDR) from an antibody generated against a Helicobacter pylori antigen  
 CC (catalase or beta-urease) which is used to illustrate the method of the  
 CC invention

XX SQ Sequence 8 AA;

Query Match 100.0%; Score 41; DB 4; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQYSSSPT 8  
 |||||  
 Db 1 QQYSSSPT 8

RESULT 3  
 AAB86108

XX ID AAB86108 standard; protein; 106 AA.

XX AC AAB86108;

XX DT 17-JUL-2001 (first entry)

XX DE H. pylori catalase antibody HP25/2m/1BS V region light chain.

XX KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;  
 acid-resistant microorganism; complementarity determining region; CDR;  
 KW feces; heavy chain; light chain.

XX OS Unidentified.

XX PN WO200127612-A2.

XX PD 19-APR-2001.

XX PF 12-OCT-2000; 2000WO-EP010057.

XX PR 12-OCT-1999; 99EP-00120351.

XX PR 16-MAR-2000; 2000EP-00105592.

XX PR 31-MAR-2000; 2000EP-00107028.

XX PR 10-MAY-2000; 2000EP-00110110.

XX PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX PI Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;

XX DR WPI; 2001-282086/29.

XX DR N-PSDB; AAF88154.

XX PT Detecting infections by acid-resistant microorganisms, particularly for  
 diagnosing Helicobacter pylori, comprises immunochromatographic detection  
 of antigen in feces.

XX PS Claim 31; Fig 4; 90pp; German.

XX This invention describes a novel method for detecting infection by an  
 CC acid-resistant microorganism (A), in a mammal, using  
 CC immunochromatography. The method is used to diagnose infection by an acid  
 CC -resistant microorganism (A), in a mammal, such as Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,  
 CC inexpensive and non-invasive, and may indicate the stage of infection. A  
 CC test strip used in the method may include a filter to eliminate particles  
 CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence represents a

CC Helicobacter pylori catalase derived antibody V-region light chain  
CC fragment used to illustrate the method of the invention  
XX  
SQ Sequence 106 AA;

Query Match 100.0%; Score 41; DB 4; Length 106;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQYSSSPT 8  
|||  
Db 89 QQYSSSPT 96

RESULT 4  
AAB86070  
ID AAB86070 standard; protein; 106 AA.  
XX  
AC AAB86070;  
XX  
DT 17-JUL-2001 (first entry)  
XX  
DE H. pylori catalase antibody HP25/6m/1B5 V-region light chain.  
XX  
KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
KW antibacterial; complementarity determining region; V-region.  
XX  
OS Unidentified.  
XX  
PN WO200127613-A2.  
XX  
PD 19-APR-2001.  
XX  
PF 12-OCT-2000; 2000WO-EP010058.  
XX  
PR 12-OCT-1999; 99EP-00120351.  
PR 16-MAR-2000; 2000EP-00105592.  
PR 31-MAR-2000; 2000EP-00107028.  
PR 10-MAY-2000; 2000EP-00110110.  
XX  
PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
XX  
PI Reiter C, Cullmann G, Heppner P, Ringeis A, Mueller H, Haindl E;  
XX  
DR WPI; 2001-282087/29.  
DR N-PSDB; AAF88097.  
XX  
PT Detecting infections by acid-resistant microorganisms, particularly for  
PT diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
PT sample.  
XX  
PS Claim 27; Fig 4; 89pp; German.  
XX  
SS This invention describes a novel method for detecting, in a mammal,  
CC infection by an acid-resistant microorganism (A) which comprises reacting  
CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
CC complex is formed with Ag, and the formation of a complex detected. R are  
CC specific for an Ag which, after passage through the intestines, at least  
CC in some mammals, retains a native (or corresponding) structure against  
CC which the mammal produces antibodies (when immunized or infected with  
CC (A) or its extracts, lysates or derived proteins (or fragments) or  
CC synthetic peptides). The products of the invention have antibacterial  
CC activity. The method is used to diagnose infection by Helicobacter,  
CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the  
CC progress of treatment. Receptors, particularly antibodies, directed  
CC against Ag can be used therapeutically for treatment of infections. The  
CC method requires only one R to provide a reasonably secure diagnosis  
CC (although use of two R improves sensitivity), so is relatively  
CC inexpensive and more easily standardized. Also it is direct, non-  
CC invasive, suitable for automation and may indicate the stage of an

CC infection. This sequence represents a Helicobacter pylori anti-catalase  
CC derived antibody HP25/6m/1B5 V-region light chain fragment which is  
CC described in the method of the invention  
XX  
SQ Sequence 106 AA;

Query Match 100.0%; Score 41; DB 4; Length 106;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQYSSSPT 8  
|||  
Db 89 QQYSSSPT 96

RESULT 5  
ABP62398  
ID ABP62398 standard; peptide; 8 AA.  
XX  
AC ABP62398;  
XX  
DT 10-OCT-2002 (first entry)  
XX  
DE Human immunopeptide to HCV E2 glycoprotein light chain CDR #75.  
XX  
KW Virucide; human; immunopolypeptide; immunopeptide; envelope glycoprotein;  
KW nonstructural protein; hepatitis C virus; HCV; E2 glycoprotein;  
KW NS3 protein; viral infection.  
XX  
OS Homo sapiens.  
XX  
PN WO200259340-A1.  
XX  
PD 01-AUG-2002.  
XX  
PF 25-JAN-2002; 2002WO-US002303.  
XX  
PR 26-JAN-2001; 2001US-0264451P.  
XX  
PA (SCRI ) SCRIPPS RES INST.  
XX  
PI Maruyama T, Jones IM, Burton DR, Fox RI;  
XX  
DR WPI; 2002-599801/64.  
XX  
PT New human immunopolypeptide with binding specificity for certain envelope  
PT glycoproteins and nonstructural proteins of hepatitis C virus (HCV), for  
PT diagnosing or treating patients having or suspected of having HCV  
PT infection.  
XX  
PS Claim 1; Fig 17; 308pp; English.  
XX  
SS The present invention relates to human immunopolypeptides, produced by a  
CC phage transfected cell library. The present sequence is one such  
CC immunopolypeptide. The immunopolypeptides have binding specificity for  
CC envelope glycoprotein E2 and nonstructural protein NS3 of hepatitis C  
CC virus (HCV). E2 glycoprotein is believed to be responsible for target  
CC cell binding and contains neutralising epitopes, while NS3 is thought to  
CC be involved in the replication of HCV. The immunopolypeptides are useful  
CC for diagnosing and treating a patient having or suspected to be having  
CC HCV infection  
XX  
SQ Sequence 8 AA;

Query Match 90.2%; Score 37; DB 5; Length 8;  
Best Local Similarity 87.5%; Pred. No. 1.8e+06;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQYSSSPT 8  
|||  
Db 1 QQYSSSPT 8

RESULT 6  
ABB07229  
ID ABB07229 standard; protein; 107 AA.  
XX  
AC ABB07229;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Anti-IL-4 receptor MAb 6-2 light chain variable region.  
XX  
KW Human; antibody; interleukin; IL-4; antagonist; receptor; IL-4 receptor;  
XX antiarthritic; dermatological; antiulcer; antiinflammatory; cytostatic;  
KW antischlicking; immunosuppressive; tuberculostatic; ophthalmological;  
KW antianemic; antithyroid.  
XX  
OS Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT Region 24..35  
FT /note= "complementarity determining region (CDR) 1"  
FT Region 51..57  
FT /note= "complementarity determining region (CDR) 2"  
FT Region 90..97  
FT /note= "complementarity determining region (CDR) 3"  
XX  
WO200192340-A2.  
XX  
PN  
XX  
PD 06-DEC-2001.  
XX  
XX  
PF 25-MAY-2001; 2001WO-US017094.  
XX  
PR 26-MAY-2000; 2000US-00579808.  
XX  
PR 19-SEP-2000; 2000US-00665343.  
XX  
PR 15-FEB-2001; 2001US-00785934.  
XX  
PR 01-MAY-2001; 2001US-00847816.  
XX  
XX (IMV ) IMMUNEX CORP.  
XX  
XX Plueneke JD;  
XX  
XX WPI; 2002-114332/15.  
XX  
DR N-PSDB; ABA94329.  
XX  
XX Novel human antibody which binds human interleukin (IL)-4 receptor and is  
XX capable of inhibiting IL-4 induced biological activity, functions as IL-4  
XX antagonist and is useful for treating septic arthritis, scleroderma.  
XX  
XX Claim 3; Page 72; 85pp; English.  
XX  
XX The invention relates to a human antibody (an interleukin (IL)-4  
XX antagonist) (I) that binds human IL-4 receptor (IL-4R), and is capable of  
XX inhibiting an IL-4-induced biological activity. (I) is also useful for  
XX inhibiting both IL-4-induced biological activity and IL-13-induced  
XX biological activity in vivo in a human, and for treating septic arthritis  
XX in a human afflicted with septic arthritis. (I) is also used for treating  
XX conditions such as septic/reactive arthritis, dermatitis herpetiformis,  
XX urticaria (especially chronic idiopathic urticaria), ulcers, gastric  
XX inflammation, mucosal inflammation, ulcerative colitis, Crohn's disease,  
XX inflammatory bowel disease, other disorders of the digestive system in  
XX which IL-4 plays a role (e.g. IL-4-induced inflammation of part of the  
XX gastrointestinal tract), conditions in which IL-4-induced barrier  
XX disruption plays a role (e.g. conditions characterized by decreased  
XX epithelial barrier function in the lung or gastrointestinal tract),  
XX scleroderma, hypertrophic scarring, Whipple's disease, benign prostate  
XX hyperplasia, IL-4-induced pulmonary conditions, allergic reactions to  
XX medication, Kawasaki disease, sickle cell disease or crisis, Churg-  
XX Strauss syndrome, Grave's disease, pre-eclampsia, Sjogren's syndrome,  
XX autoimmune lymphoproliferative syndrome, autoimmune haemolytic anemia,  
XX Barrett's esophagus, autoimmune uveitis, tuberculosis, nephrosis,  
XX pemphigus vulgaris or bullous pemphigoid (autoimmune blistering  
XX disease), and myasthenia gravis (an autoimmune muscular disease). IL-4  
XX antagonists also find use as adjuvants to allergy immunotherapy and as  
XX vaccine adjuvants, especially when directing the immune response toward a

CC TH1 response would be beneficial in treating or preventing the disease.  
CC The present sequence represents an anti-IL-4 receptor monoclonal antibody  
CC (MAb) 6-2 light chain variable region  
XX  
SQ Sequence 107 AA;  
XX  
Query Match 90.2%; Score 37; DB 5; Length 107;  
Best Local Similarity 87.5%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 QQYSSSPT 8  
DB 90 QQYSSSPT 97  
XX  
RESULT 7  
AAB62755  
ID AAB62755 standard; protein; 108 AA.  
XX  
AC AAB62755;  
XX  
DT 03-APR-2001 (first entry)  
XX  
DE Human HIV-1 monoclonal antibody SEQ ID NO: 54.  
XX  
KW Human immunodeficiency virus-1; HIV-1; human monoclonal antibody;  
XX envelope glycoprotein; gp120; diagnosis.  
XX  
OS Homo sapiens.  
XX  
PN WO200100678-A1.  
XX  
PD 04-JAN-2001.  
XX  
PF 23-JUN-2000; 2000WO-US017327.  
XX  
PR 30-JUN-1999; 99US-0141701P.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Watkins BA, Reitz MS;  
XX  
XX WPI; 2001-112438/12.  
XX  
DR N-PSDB; AAF29056.  
XX  
XX Novel human monoclonal antibody immunoreactive with human  
XX immunodeficiency virus-1 glycoprotein gp120, useful for detecting HIV-1  
XX in biological sample and providing passive immunotherapy to HIV-1  
XX infected mammal.  
XX  
XX Claim 1; Page 56-57; 81pp; English.  
XX  
XX The present invention provides the protein and coding sequences for the  
XX variable regions of human monoclonal antibodies which are immunoreactive  
XX with human immunodeficiency virus-1 (HIV-1) envelope glycoprotein gp120.  
XX These can be used in diagnosis and therapy of HIV-1 infection  
XX  
XX Sequence 108 AA;  
XX  
Query Match 90.2%; Score 37; DB 4; Length 108;  
Best Local Similarity 87.5%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 QQYSSSPT 8  
DB 90 QQYSSSPT 97  
XX  
RESULT 8  
ABP62189  
ID ABP62189 standard; protein; 132 AA.  
XX  
AC ABP62189;

XX 10-OCT-2002 (first entry)  
 DT Human immunopeptide to HCV E2 glycoprotein Fab variable region #36.  
 DE  
 XX Virucide; human; immunopolypeptide; immunopeptide; envelope glycoprotein;  
 KW nonstructural protein; hepatitis C virus; HCV; E2 glycoprotein;  
 KW NS3 protein; viral infection.  
 XX  
 OS Homo sapiens.  
 XX WO200259340-A1.  
 PN  
 XX 01-AUG-2002.  
 PD  
 XX 25-JAN-2002; 2002WO-US0002303.  
 XX  
 PF 26-JAN-2001; 2001US-0264451P.  
 XX  
 PR (SCRI ) SCRIPPS RES INST.  
 XX  
 PA Maruyama T, Jones IM, Burton DR, Fox RI;  
 PI WPI; 2002-599801/64.  
 XX N-PSDB; AB092688.  
 DR  
 XX New human immunopolypeptide with binding specificity for certain envelope  
 PT glycoproteins and nonstructural proteins of hepatitis C virus (HCV), for  
 PT diagnosing or treating patients having or suspected of having HCV  
 PT infection.  
 XX  
 XX Claim 10; Fig 12; 308pp; English.  
 PS  
 XX The present invention relates to human immunopolypeptides, produced by a  
 CC phage transfected cell library. The present sequence is one such  
 CC immunopolypeptide. The immunopolypeptides have binding specificity for  
 CC envelope glycoprotein E2 and nonstructural protein NS3 of hepatitis C  
 CC virus (HCV). E2 glycoprotein is believed to be responsible for target  
 CC cell binding and contains neutralising epitopes, while NS3 is thought to  
 CC be involved in the replication of HCV. The immunopolypeptides are useful  
 CC for diagnosing and treating a patient having or suspected to be having  
 CC HCV infection  
 XX  
 XX Sequence 132 AA;  
 SQ  
 Query Match 90.2%; Score 37; DB 5; Length 132;  
 Best Local Similarity 87.5%; Pred. No. 16;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QQYSSSPT 8  
 DB 90 QQYSSSPT 97  
 RESULT 9  
 AAM24101  
 ID AAM24101 standard; protein; 384 AA.  
 XX  
 AC AAM24101;  
 DT  
 XX 12-OCT-2001 (first entry)  
 DT  
 XX Human EST encoded protein SEQ ID NO: 1626.  
 DE  
 XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
 KW diagnostics; forensic test; Gene mapping; genetic disorder; biodiversity;  
 KW gene therapy; nutrition.  
 XX  
 OS Homo sapiens.  
 XX WO200154477-A2.  
 PN  
 XX

PD 02-AUG-2001.  
 XX  
 PF 25-JAN-2001; 2001WO-US002687.  
 XX  
 PR 25-JAN-2000; 2000US-00491404.  
 PR 17-JUL-2000; 2000US-00617746.  
 PR 03-AUG-2000; 2000US-00631451.  
 PR 15-SEP-2000; 2000US-00663870.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
 PI WPI; 2001-476164/51.  
 XX N-PSDB; AAH98760.  
 DR  
 XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
 PT antibodies and research use.  
 PT  
 XX Claim 20; Page 1102-1103; 1275pp; English.  
 PS  
 XX The present invention provides the protein and coding sequences of novel  
 CC proteins from a variety of organisms, including human, dog, cat, horse,  
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
 CC from the organism of interest. They can be used in diagnostics,  
 CC forensics, gene mapping, identification of mutations, to assess  
 CC biodiversity and for nutritional purposes. The present sequence is a  
 CC protein of the invention  
 XX  
 XX Sequence 384 AA;  
 SQ  
 Query Match 90.2%; Score 37; DB 4; Length 384;  
 Best Local Similarity 87.5%; Pred. No. 53;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QQYSSSPT 8  
 DB 256 QQYSSSPT 263  
 RESULT 10  
 AAU14462  
 ID AAU14462 standard; protein; 384 AA.  
 XX  
 AC AAU14462;  
 XX  
 DT 24-OCT-2001 (first entry)  
 DT  
 XX Human novel protein #333.  
 DE  
 XX Human; novel protein; Antianaemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cytostatic; neuroprotective; vulnery; nootropic;  
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200155437-A2.  
 PN  
 XX 02-AUG-2001.  
 PD  
 XX 25-JAN-2001; 2001WO-US002623.  
 PF  
 XX 25-JAN-2000; 2000US-00491404.  
 PR  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Tang YT, Liu C, Drmanac RT;  
 PI

```

XX WPI; 2001-451939/48.
DR N-PSDB; AAS22767.
XX
PT Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage.
XX
PS Example 4; Page 825-826; 894pp; English.
XX
CC The invention relates to polynucleotides encoding novel human proteins or
CC their active domains. The polypeptides, polynucleotides and antibodies
CC raised against the polypeptides are used in a method of treatment of a
CC mammal and prevention of disorders caused by the aberrant protein
CC expression or activity. The polypeptides can be used as molecular weight
CC markers, food supplements, and in antibody production. The polypeptides
CC are used to identify compounds which bind to the polypeptides.
CC Polynucleotides of the invention are used as probes and primers, for
CC sequencing, for chromosome or gene mapping, in the production of
CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene
CC therapy. Polypeptides of the invention can be used to target drugs to a
CC tumour, in assays to determine biological activity, to raise
CC antibodies/elicit an immune response, to determine quantitative protein
CC levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
CC diseases, nervous system disorders, and infection. The present sequence
CC represents a protein of the invention
XX
SQ Sequence 384 AA;
Query Match 90.2%; Score 37; DB 4; Length 384;
Best Local Similarity 87.5%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQYSSSPT 8
Db 259 QQYGSSTP 266
|||||

RESULT 11
AAU14463
ID AAU14463 standard; protein; 384 AA.
XX
AC AAU14463;
XX
XX 24-OCT-2001 (first entry)
XX
DE Human novel protein #334.
XX
KW Human; novel protein; Antianaemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytosstatic; neuroprotective; vulnery; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
XX WO200155437-A2.
XX
PD 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US002623.
XX
XX 25-JAN-2000; 2000US-00491404.
PR

```

```

XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-451939/48.
DR N-PSDB; AAS22768.
XX
PT Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage.
XX
PS Example 4; Page 826-827; 894pp; English.
XX
CC The invention relates to polynucleotides encoding novel human proteins or
CC their active domains. The polypeptides, polynucleotides and antibodies
CC raised against the polypeptides are used in a method of treatment of a
CC mammal and prevention of disorders caused by the aberrant protein
CC expression or activity. The polypeptides can be used as molecular weight
CC markers, food supplements, and in antibody production. The polypeptides
CC are used to identify compounds which bind to the polypeptides.
CC Polynucleotides of the invention are used as probes and primers, for
CC sequencing, for chromosome or gene mapping, in the production of
CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene
CC therapy. Polypeptides of the invention can be used to target drugs to a
CC tumour, in assays to determine biological activity, to raise
CC antibodies/elicit an immune response, to determine quantitative protein
CC levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
CC diseases, nervous system disorders, and infection. The present sequence
CC represents a protein of the invention
XX
SQ Sequence 384 AA;
Query Match 90.2%; Score 37; DB 4; Length 384;
Best Local Similarity 87.5%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQYSSSPT 8
Db 259 QQYGSSTP 266
|||||

RESULT 12
AAU14461
ID AAU14461 standard; protein; 384 AA.
XX
AC AAU14461;
XX
XX 24-OCT-2001 (first entry)
XX
DE Human novel protein #332.
XX
KW Human; novel protein; Antianaemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytosstatic; neuroprotective; vulnery; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
XX WO200155437-A2.
XX
XX 02-AUG-2001.
PD

```



```

XX PF 25-JAN-2001; 2001WO-US0002623.
XX PR 25-JAN-2000; 2000US-00491404.
XX PR (HYSE-) HYSEQ INC.
XX PA Tang YT, Liu C, Drmanac RT;
XX PI WPI; 2001-451939/48.
XX DR N-PSDB; AAS22766.
XX PT Isolated polypeptides useful for treating anti-inflammatory diseases,
XX PT nervous system disorders, and for regenerating bone and cartilage.
XX PS Example 4; Page 824-825; 894pp; English.
XX CC The invention relates to polynucleotides encoding novel human proteins or
XX CC their active domains. The polypeptides, polynucleotides and antibodies
XX CC raised against the polypeptides are used in a method of treatment of a
XX CC mammal and prevention of disorders caused by the aberrant protein
XX CC expression or activity. The polypeptides can be used as molecular weight
XX CC markers, food supplements, and in antibody production. The polypeptides
XX CC are used to identify compounds which bind to the polypeptides.
XX CC Polynucleotides of the invention are used as probes and primers, for
XX CC sequencing, for chromosome or gene mapping, in the production of
XX CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene
XX CC therapy. Polypeptides of the invention may also be used to target drugs to a
XX CC tumour, in assays to determine biological activity, to raise
XX CC antibodies/elicite an immune response, to determine quantitative protein
XX CC levels, as tissue markers, and to isolate receptors or ligands.
XX CC Polypeptides of the invention may also be useful in treating platelet
XX CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
XX CC ligament and/or nerve tissue, wound healing, treating burns, promoting
XX CC the proliferation, differentiation and survival of stem cells, as a
XX CC contractile, treating osteoporosis and osteoarthritis, anaemia,
XX CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
XX CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
XX CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
XX CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
XX CC diseases, nervous system disorders, and infection. The present sequence
XX CC represents a protein of the invention
XX SQ Sequence 384 AA;

Query Match 90.2%; Score 37; DB 4; Length 384;
Best Local Similarity 87.5%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQYSSSPT 8
Db 259 QQYSSSPT 266

RESULT 13
AAU14464
XX ID AAU14464 standard; protein; 384 AA.
XX AC AAU14464;
XX DT 24-OCT-2001 (first entry)
XX DE Human novel protein #335.
XX KW Human, novel protein; Antianemic; osteopathic; antiinflammatory;
XX KW immunomodulatory; cytostatic; neuroprotective; vulnary; nootropic;
XX KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
XX KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
XX KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
XX KW Parkinson's disease; inflammatory disorder; cancer; asthma, osteoporosis;
XX KW tissue regeneration; immune disorder.
XX OS Homo sapiens.

XX PN WO200155437-A2.
XX XX 02-AUG-2001.
XX PF 25-JAN-2001; 2001WO-US0002623.
XX PR 25-JAN-2000; 2000US-00491404.
XX PR (HYSE-) HYSEQ INC.
XX PA Tang YT, Liu C, Drmanac RT;
XX PI WPI; 2001-451939/48.
XX DR N-PSDB; AAS22769.
XX PT Isolated polypeptides useful for treating anti-inflammatory diseases,
XX PT nervous system disorders, and for regenerating bone and cartilage.
XX PS Example 4; Page 827; 894pp; English.
XX CC The invention relates to polynucleotides encoding novel human proteins or
XX CC their active domains. The polypeptides, polynucleotides and antibodies
XX CC raised against the polypeptides are used in a method of treatment of a
XX CC mammal and prevention of disorders caused by the aberrant protein
XX CC expression or activity. The polypeptides can be used as molecular weight
XX CC markers, food supplements, and in antibody production. The polypeptides
XX CC are used to identify compounds which bind to the polypeptides.
XX CC Polynucleotides of the invention are used as probes and primers, for
XX CC sequencing, for chromosome or gene mapping, in the production of
XX CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene
XX CC therapy. Polypeptides of the invention may also be used to target drugs to a
XX CC tumour, in assays to determine biological activity, to raise
XX CC antibodies/elicite an immune response, to determine quantitative protein
XX CC levels, as tissue markers, and to isolate receptors or ligands.
XX CC Polypeptides of the invention may also be useful in treating platelet
XX CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
XX CC ligament and/or nerve tissue, wound healing, treating burns, promoting
XX CC the proliferation, differentiation and survival of stem cells, as a
XX CC contractile, treating osteoporosis and osteoarthritis, anaemia,
XX CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
XX CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
XX CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
XX CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
XX CC diseases, nervous system disorders, and infection. The present sequence
XX CC represents a protein of the invention
XX SQ Sequence 384 AA;

Query Match 90.2%; Score 37; DB 4; Length 384;
Best Local Similarity 87.5%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQYSSSPT 8
Db 259 QQYSSSPT 266

RESULT 14
ADH80782
XX ID ADH80782 standard; protein; 384 AA.
XX AC ADH80782;
XX DT 22-APR-2004 (first entry)
XX DE Human polypeptide #99.
XX KW Human, coagulation disorder; haemophilia; wound; stroke; thrombosis;
XX KW myocardial infarction; cancer; bone fracture; Alzheimer's disease;
XX KW Parkinson's disease; autoimmune disorder; food supplement; haemostatic;
XX KW vulnary; cerebroprotective; thrombolytic; anticoagulant; cardiant;
XX KW cyostatic; osteopathic; neuroprotective; nootropic; antiparkinsonian;

```

KW immunosuppressive.

OS Homo sapiens.

XX US2003232054-A1.

XX 18-DEC-2003.

XX 08-NOV-2002; 2002US-00291265.

XX 25-JAN-2000; 2000US-00491404.

XX 17-JUL-2000; 2000US-00617746.

PR 03-AUG-2000; 2000US-00631451.

PR 15-SEP-2000; 2000US-00663870.

PR 25-JAN-2001; 2001WO-US002623.

PR 03-AUG-2001; 2001US-00922279.

XX (TANG/) TANG Y T.

PA (LIUC/) LIU C.

PA (ASUN/) ASUNDI V.

PA (CHEN/) CHEN R.

PA (QIAN/) QIAN X B.

PA (WANG/) WANG Z W.

PA (WEHR/) WEHRMAN T.

PA (ZHAN/) ZHANG J.

PA (ZHOU/) ZHOU P.

PA (CAOY/) CAO Y.

PA (DRMA/) DRMANAC R T.

XX Tang YT, Liu C, Asundi V, Chen R, Qian XB, Wang ZW, Wehrman T;

PI Zhang J, Zhou P, Cao Y, Drmanac RT;

XX WPI; 2004-061257/06.

XX New polynucleotides and polypeptides useful for diagnosing, preventing or

PT treating diseases involving aberrant protein expression or activity, e.g.

PT hemophilia, wounds, stroke, thrombosis, cancer or autoimmune disorders.

XX Claim 20; SEQ ID NO 807; 85pp; English.

XX The invention relates to new isolated polynucleotides and polypeptides.

CC The sequences, compositions and methods of the invention are useful for

CC diagnosing, preventing or treating diseases involving aberrant protein

CC expression or biological activity, such as coagulation disorders (e.g.

CC haemophilia), wounds, stroke, thrombosis, myocardial infarction, cancer,

CC bone fractures, Alzheimer's disease, Parkinson's disease and autoimmune

CC disorders. The polynucleotides may be used as hybridisation probes, as

CC oligomers or primers, for polymerase chain reaction, for chromosome and

CC gene mapping, in the recombinant production of proteins and in generation

CC of antisense DNA or RNA. The polypeptides may be used in generating

CC antibodies, as molecular weight markers or as food supplements. This

XX sequence represents a human polypeptide of the invention.

XX Query Match

XX Best Local Similarity

XX Matches

XX 7; Conservative

XX 0; Mismatches

XX 1; Indels

XX 0; Gaps

XX 0;

QY 1 QQYSSSPT 8

Db 259 QQYGSST 266

RESULT 15

ADH80779

ID ADH80779 standard; protein; 384 AA.

XX AC

XX ADH80779;

XX 22-APR-2004 (first entry)

XX Human polypeptide #96.

DE

XX

KW

KW

KW

KW

KW

KW

KW

XX

OS

XX

XX

PN

XX

PD

XX

PF

XX

PR

PR

PR

PR

PR

PR

PR

XX

PA

PA

PA

PA

PA

PA

PA

PA

PA

XX

PI

PI

XX

DR

Human; coagulation disorder; haemophilia; wound; stroke; thrombosis;

myocardial infarction; cancer; bone fracture; Alzheimer's disease;

Parkinson's disease; autoimmune disorder; food supplement; haemostatic;

vulnary; cerebroprotective; thrombolytic; anticoagulant; cardiant;

cytostatic; osteopathic; neuroprotective; nootropic; antiparkinsonian;

immunosuppressive.

Homo sapiens.

US2003232054-A1.

18-DEC-2003.

08-NOV-2002; 2002US-00291265.

25-JAN-2000; 2000US-00491404.

17-JUL-2000; 2000US-00617746.

03-AUG-2000; 2000US-00631451.

15-SEP-2000; 2000US-00663870.

25-JAN-2001; 2001WO-US002623.

03-AUG-2001; 2001US-00922279.

(TANG/) TANG Y T.

(LIUC/) LIU C.

(ASUN/) ASUNDI V.

(CHEN/) CHEN R.

(QIAN/) QIAN X B.

(WANG/) WANG Z W.

(WEHR/) WEHRMAN T.

(ZHAN/) ZHANG J.

(ZHOU/) ZHOU P.

(CAOY/) CAO Y.

(DRMA/) DRMANAC R T.

Tang YT, Liu C, Asundi V, Chen R, Qian XB, Wang ZW, Wehrman T;

Zhang J, Zhou P, Cao Y, Drmanac RT;

WPI; 2004-061257/06.

New polynucleotides and polypeptides useful for diagnosing, preventing or

treating diseases involving aberrant protein expression or activity, e.g.

hemophilia, wounds, stroke, thrombosis, cancer or autoimmune disorders.

Claim 20; SEQ ID NO 804; 85pp; English.

The invention relates to new isolated polynucleotides and polypeptides.

The sequences, compositions and methods of the invention are useful for

diagnosing, preventing or treating diseases involving aberrant protein

expression or biological activity, such as coagulation disorders (e.g.

haemophilia), wounds, stroke, thrombosis, myocardial infarction, cancer,

bone fractures, Alzheimer's disease, Parkinson's disease and autoimmune

disorders. The polynucleotides may be used as hybridisation probes, as

oligomers or primers, for polymerase chain reaction, for chromosome and

gene mapping, in the recombinant production of proteins and in generation

of antisense DNA or RNA. The polypeptides may be used in generating

antibodies, as molecular weight markers or as food supplements. This

sequence represents a human polypeptide of the invention.

Query Match

Best Local Similarity

Matches

7; Conservative

0; Mismatches

1; Indels

0; Gaps

0;

QY 1 QQYSSSPT 8

Db 259 QQYGSST 266

Search completed: August 9, 2005, 14:14:58

Job time : 102.538 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:07:04 ; Search time 24 Seconds  
(without alignments)  
24.883 Million cell updates/sec

Title: US-10-089-452-29  
Perfect score: 41  
Sequence: 1 QYSSSPT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	80.5	153	4	US-09-270-767-3335, A
2	33	80.5	153	4	US-09-270-767-48552, A
3	33	80.5	715	4	US-09-248-796A-20602
4	33	80.5	1210	2	US-08-484-438-7
5	33	80.5	1210	2	US-08-475-035-4
6	33	80.5	1210	4	US-09-715-249-2
7	32	78.0	8	4	US-09-563-222C-16
8	32	78.0	9	2	US-08-350-260A-503
9	32	78.0	9	3	US-09-042-353-362
10	32	78.0	9	3	US-08-758-417A-210
11	32	78.0	9	4	US-09-104-337A-503
12	32	78.0	10	4	US-09-563-222C-13
13	32	78.0	96	4	US-09-472-087-87
14	32	78.0	107	1	US-08-107-669D-14
15	32	78.0	107	1	US-08-472-788A-14
16	32	78.0	107	2	US-08-477-531B-14
17	32	78.0	107	2	US-08-652-558-49
18	32	78.0	107	2	US-08-082-842A-14
19	32	78.0	107	4	US-08-635-109-7
20	32	78.0	107	4	US-08-844-215-10
21	32	78.0	108	1	US-08-488-113B-150
22	32	78.0	108	1	US-08-477-484B-150
23	32	78.0	108	2	US-08-646-360-150
24	32	78.0	108	2	US-08-232-081B-42
25	32	78.0	108	3	US-08-839-765-150
26	32	78.0	108	3	US-09-136-389-150
27	32	78.0	108	3	US-09-240-274-178

28	32	78.0	108	3	US-09-610-838-150	Sequence 150, App
29	32	78.0	108	4	US-09-711-485-150	Sequence 150, App
30	32	78.0	112	1	US-08-276-852-151	Sequence 151, App
31	32	78.0	112	1	US-08-899-575-151	Sequence 151, App
32	32	78.0	112	1	US-08-899-575-151	Sequence 151, App
33	32	78.0	112	5	PCT-US95-08743-151	Sequence 151, App
34	32	78.0	116	1	US-08-053-131-183	Sequence 183, App
35	32	78.0	116	2	US-08-096-762-183	Sequence 183, App
36	32	78.0	116	3	US-09-042-353-46	Sequence 183, App
37	32	78.0	116	3	US-08-758-417A-311	Sequence 311, App
38	32	78.0	150	3	US-08-862-124-5	Sequence 5, Appl
39	32	78.0	226	4	US-09-456-090A-38	Sequence 38, Appl
40	32	78.0	226	4	US-09-456-090A-42	Sequence 42, Appl
41	32	78.0	226	4	US-09-456-090A-50	Sequence 50, Appl
42	32	78.0	226	4	US-09-456-090A-72	Sequence 72, Appl
43	32	78.0	226	4	US-09-456-090A-74	Sequence 74, Appl
44	32	78.0	226	4	US-09-456-090A-80	Sequence 80, Appl
45	32	78.0	226	4	US-09-456-090A-86	Sequence 86, Appl

ALIGNMENTS

RESULT 1  
US-09-270-767-3335  
; Sequence 3335, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3335  
; LENGTH: 153  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-3335

Query Match 80.5%; Score 33; DB 4; Length 153;  
Best Local Similarity 85.7%; Pred. No. 39;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYSSSPT 8  
Db 126 QYTSPT 132

RESULT 2  
US-09-270-767-48552  
; Sequence 48552, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48552  
; LENGTH: 153  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-48552

Query Match 80.5%; Score 33; DB 4; Length 153;  
Best Local Similarity 85.7%; Pred. No. 39;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYSSSPT 8

```
Db      126 QYSSPT 132
||:||||
RESULT 3
US-09-248-796A-20602
; Sequence 20602, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20602
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20602
Query Match      80.5%; Score 33; DB 4; Length 715;
Best Local Similarity 87.5%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 QYSSPT 8
||:||||
Db      596 QQSSSPT 603
||:||||
RESULT 4
US-08-484-438-7
; Sequence 7, Application US/08484438
; Patent No. 5811098
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Culouscou, Jean-Michel
; APPLICANT: Shoyab, Mohammed
; APPLICANT: Siegall, Clay B.
; APPLICANT: Hellstr m, Ingegerd
; APPLICANT: Hellstr m, Karl E.
; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,438
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,442
; FILING DATE: 14-OCT-1994
; APPLICATION NUMBER: US 08/150,704
; FILING DATE: 10-NOV-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/981,165
```

```
; FILING DATE: 24-NOV-1992
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leelle
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 5624-230
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1210 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-484-438-7
Query Match      80.5%; Score 33; DB 2; Length 1210;
Best Local Similarity 75.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 QYSSPT 8
||:||||
Db      1067 QYSSDPT 1074
||:||||
RESULT 5
US-08-475-035-4
; Sequence 4, Application US/08475035
; Patent No. 5985553
; GENERAL INFORMATION:
; APPLICANT: KING, C. R.
; APPLICANT: KRAUS, MATTHIAS H.
; TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM
; TITLE OF INVENTION: EGF RECEPTOR GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, 127 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,035
; FILING DATE: 7 Jun 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.656
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1210 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-475-035-4
Query Match      80.5%; Score 33; DB 2; Length 1210;
Best Local Similarity 75.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 QOYSSSPT 8
Db      1067 QRYSDPT 1074

RESULT 6
US-09-715-249-2
; Sequence 2, Application US/09715249
; Patent No. 6790614
; GENERAL INFORMATION:
; APPLICANT: NOVARTIS AG
; APPLICANT: VERES, GABOR
; APPLICANT: PIPPIG, SUSANNE
; TITLE OF INVENTION: selectable cell surface marker genes
; FILE REFERENCE: 4-31192
; CURRENT APPLICATION NUMBER: US/09/715,249
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: us 60/166594
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: us 09/539248
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: EGFR
US-09-715-249-2

Query Match      80.5%; Score 33; DB 4; Length 1210;
Best Local Similarity 75.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 QOYSSSPT 8
Db      1067 QRYSDPT 1074

RESULT 7
US-09-563-222C-16
; Sequence 16, Application US/09563222C
; Patent No. 6696620
; GENERAL INFORMATION:
; APPLICANT: EPICYTE PHARMACEUTICALS, INC.
; APPLICANT: HIATT, ANDREW C.
; APPLICANT: HEIN, MICH B.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN PLANT CELLS
; FILE REFERENCE: 068904-0501
; CURRENT APPLICATION NUMBER: US/09/563,222C
; CURRENT FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: PCT/US01/14349
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 09/563,222
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-563-222C-16

Query Match      78.0%; Score 32; DB 4; Length 8;
Best Local Similarity 75.0%; Pred. No. 4.1e+05;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 QOYSSSPT 8
Db      1 QOYYSPT 8

RESULT 8
US-08-350-260A-503
; Sequence 503, Application US/08350260A
; Patent No. 5962255
; GENERAL INFORMATION:
; APPLICANT: Winter, Gregory Paul
; APPLICANT: Griffiths, Andrew David
; APPLICANT: Williams, Samuel Cameron
; APPLICANT: Waterhouse, Peter
; APPLICANT: Nissim, Ahuva
; APPLICANT: Johnson, Kevin Stuart
; APPLICANT: Smith, Andrew John Hammond
; TITLE OF INVENTION: Methods for producing members of specific
; TITLE OF INVENTION: binding pairs
; NUMBER OF SEQUENCES: 602
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David W. Clough
; STREET: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/08/350,260A
; FILING DATE: 05-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9110549.4
; FILING DATE: 15-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206318.9
; FILING DATE: 24-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB91/01134
; FILING DATE: 10-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/00883
; FILING DATE: 15-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00605
; FILING DATE: 24-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/150,002
; FILING DATE: 31-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/307,619
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28111/32372
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; INFORMATION FOR SEQ ID NO: 503:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-350-260A-503

Query Match      78.0%; Score 32; DB 2; Length 9;
Best Local Similarity 85.7%; Pred. No. 4.1e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 QOYSSSPT 7
Db      1 QOYGSSP 7
```



```

; APPLICATION NUMBER: US 07/990,860
; FILING DATE: 16-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Serafini, Andrew T.
; REGISTRATION NUMBER: 41,303
; REFERENCE/DOCKET NUMBER: 014643-009030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 210:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 210:
US-08-758-417A-210

      Query Match      78.0%; Score 32; DB 3; Length 9;
      Best Local Similarity 85.7%; Pred. No. 4.1e+05;
      Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 QQYSSSP 7
Db      ||| |||
        1 QQYGSSP 7

RESULT 11
US-09-104-337A-503
; Sequence 503, Application US/09104337A
; Patent No. 6492160
; GENERAL INFORMATION:
; APPLICANT: Winter, Gregory Paul
; Griffiths, Andrew David
; Williams, Samuel Cameron
; Waterhouse, Peter
; Nissim, Ahuva
; Johnson, Kevin Stuart
; Smith, Andrew John Hammond
; TITLE OF INVENTION: Methods for producing members of specific
; binding pairs
; NUMBER OF SEQUENCES: 600
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Audrey L. Bartnicki
; STREET: Marshall, Gerstein & Borun
; 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,337A
; FILING DATE: 25-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/350,260
; FILING DATE: 05-DEC-1994
; APPLICATION NUMBER: GB 9110549.4
; FILING DATE: 15-MAY-1991
; APPLICATION NUMBER: GB 9206318.9
; FILING DATE: 24-MAR-1992
; APPLICATION NUMBER: PCT/GB92/00883
; FILING DATE: 15-MAY-1992
; APPLICATION NUMBER: PCT/GB93/00605
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: US 08/150,002
; FILING DATE: 31-MAR-1994
; APPLICATION NUMBER: US 08/307,619

; APPLICATION NUMBER: US 07/990,860
; FILING DATE: 16-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bartnicki, Audrey L.
; REGISTRATION NUMBER: 40,499
; REFERENCE/DOCKET NUMBER: 28111/32372A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; INFORMATION FOR SEQ ID NO: 503:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 503:
US-09-104-337A-503

      Query Match      78.0%; Score 32; DB 4; Length 9;
      Best Local Similarity 85.7%; Pred. No. 4.1e+05;
      Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 QQYSSSP 7
Db      ||| |||
        1 QQYGSSP 7

RESULT 12
US-09-563-222C-13
; Sequence 13, Application US/09563222C
; Patent No. 6696620
; GENERAL INFORMATION:
; APPLICANT: EPICTE PHARMACEUTICALS, INC.
; APPLICANT: HIATT, ANDREW C.
; APPLICANT: HEIN, MICH B.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN PLANT CELLS
; FILE REFERENCE: 068904-0501
; CURRENT APPLICATION NUMBER: US/09/563,222C
; CURRENT FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: PCT/US01/14349
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 09/563,222
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-563-222C-13

      Query Match      78.0%; Score 32; DB 4; Length 10;
      Best Local Similarity 85.7%; Pred. No. 3.4;
      Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 QQYSSSP 7
Db      ||| |||
        1 QQYGSSP 7

RESULT 13
US-09-472-087-87
; Sequence 87, Application US/09472087
; Patent No. 6682736
; GENERAL INFORMATION:
; APPLICANT: HANSON, DOUGLAS C.
; APPLICANT: NEVEU, MARK J.
; APPLICANT: MUELLER, ELLEN E.
; APPLICANT: HANKE, JEFFREY H.
; APPLICANT: GILMAN, STEVEN C.
; APPLICANT: DAVIS, C. GEOFFREY
; APPLICANT: CORVALAN, JOSE R.
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PFI
; CURRENT APPLICATION NUMBER: US/09/472,087

```

```

; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/113,647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-472-087-87

```

Query Match 78.0%; Score 32; DB 4; Length 96;  
Best Local Similarity 85.7%;  
Pred. NO. 37;  
Matches 6; Conservative 0; Mismatches 1; Indels

Qy 1 QYSSSP 7  
Dp 90 OYGSPP 96

RESULT 14  
US-08-107-669D-14  
; Sequence 14, Application US/08107669D  
; Patent No. 5766886  
; GENERAL INFORMATION:  
; APPLICANT: Studnicka, Gary M.  
; TITLE OF INVENTION: Modified Antibody Variable Domains (as amended)  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.  
; STREET: 1100 New York Ave., N.W., Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: United States of America

Query Match 78.0%; Score 32; DB 1; Length 107;  
Best Local Similarity 85.7%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 1; Indels

Qy 1 QYSSSP 7  
Db 89 QOYGSSP 95

```

RESULT 15
US-08-472-788A-14
; Sequence 14, Application US/08472788A
; Patent No. 5770196
; GENERAL INFORMATION:
; APPLICANT: Studnicka, Gary M.
; TITLE OF INVENTION: Modified Antibody Variable Domains
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,788A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/082,842
; FILING DATE: 23-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10906
; FILING DATE: 14-DEC-1992
; APPLICATION DATA: US 07/808,464
; FILING DATE: 13-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0610.1000003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/371-2600
; TELEFAX: 202/371-2540
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-472-788A-14

```

```
Query Match      78.0%; Score 32; DB 1; Length 107;
Best Local Similarity 85.7%; Pred. No.42;
Matches 6: Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 QYSSSP 7  
p'b 89 OYGSPP 95

Search completed: August 9, 2005, 14:22:07  
Job time : 25 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:09:49 ; Search time 86.4615 Seconds  
(without alignments)  
36.122 Million cell updates/sec

Title: US-10-089-452-29  
Perfect score: 41  
Sequence: 1 QQYSSSPT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1752860 seqs, 390397842 residues

Total number of hits satisfying chosen parameters: 1752860

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
  - 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
  - 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	37	90.2	384	15	US-10-291-265-804
2	37	90.2	384	15	Sequence 804, App
3	37	90.2	384	15	Sequence 805, App
4	37	90.2	384	15	Sequence 806, App
5	36	87.8	9	16	US-10-291-265-806
6	36	87.8	9	16	Sequence 807, App
7	36	87.8	9	16	US-10-714-353-22
8	36	87.8	9	16	Sequence 22, Appl
9	36	87.8	9	16	US-10-714-353-34
10	36	87.8	9	16	Sequence 34, Appl
11	36	87.8	9	16	US-10-714-353-40
12	36	87.8	9	16	Sequence 40, Appl
13	36	87.8	9	16	US-10-783-311-234
14	36	87.8	107	16	Sequence 234, App
15	36	87.8	116	17	US-10-723-434-16
16	36	87.8	120	16	Sequence 16, Appl
17	36	87.8	120	16	US-10-783-311-230
18	36	87.8	120	16	Sequence 230, App
19	36	87.8	120	16	US-10-714-353-44
20	36	87.8	120	16	Sequence 44, Appli

12	36	87.8	120	16	US-10-714-353-12	Sequence 12, Appl
13	33	80.5	9	15	US-10-365-761B-19	Sequence 19, Appl
14	33	80.5	11	18	US-10-862-195-1950	Sequence 1950, Ap
15	33	80.5	68	16	US-10-437-963-185574	Sequence 185574,
16	33	80.5	107	16	US-10-723-434-19	Sequence 19, Appl
17	33	80.5	107	16	US-10-723-434-24	Sequence 24, Appl
18	33	80.5	107	16	US-10-723-434-29	Sequence 29, Appl
19	33	80.5	113	16	US-10-425-115-249089	Sequence 249089,
20	33	80.5	251	15	US-09-880-748-314	Sequence 314, App
21	33	80.5	251	15	US-10-293-418-314	Sequence 314, App
22	33	80.5	327	16	US-10-437-963-133087	Sequence 133087,
23	33	80.5	644	14	US-10-259-451-2	Sequence 2, Appli
24	33	80.5	919	17	US-10-877-773-135	Sequence 135, App
25	33	80.5	919	17	US-10-877-774-135	Sequence 135, App
26	33	80.5	942	16	US-10-749-104-12	Sequence 12, Appl
27	33	80.5	1155	15	US-10-236-417-74	Sequence 116, App
28	33	80.5	1186	17	US-10-872-198-116	Sequence 116, App
29	33	80.5	1186	17	US-10-877-773-134	Sequence 134, App
30	33	80.5	1186	17	US-10-877-774-134	Sequence 134, App
31	33	80.5	1198	17	US-10-461-862-71	Sequence 71, Appl
32	33	80.5	1204	17	US-10-461-862-78	Sequence 78, Appl
33	33	80.5	1210	9	US-09-725-433-2	Sequence 2, Appli
34	33	80.5	1210	14	US-10-099-007A-17	Sequence 17, Appl
35	33	80.5	1210	15	US-10-394-322A-16	Sequence 16, Appl
36	33	80.5	1210	15	US-10-387-252-2	Sequence 2, Appli
37	33	80.5	1210	15	US-10-236-417-76	Sequence 76, Appl
38	33	80.5	1210	16	US-10-701-490-7	Sequence 7, Appli
39	33	80.5	1210	16	US-10-734-564-125	Sequence 125, App
40	33	80.5	1210	16	US-10-749-104-2	Sequence 2, Appli
41	33	80.5	1210	16	US-10-749-104-14	Sequence 14, Appl
42	33	80.5	1210	17	US-10-482-029-293	Sequence 293, App
43	33	80.5	1210	17	US-10-894-359-67	Sequence 67, Appl
44	33	80.5	1210	17	US-10-666-927-10	Sequence 10, Appl
45	33	80.5	1210	18	US-10-712-124-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1  
US-10-291-265-804  
; Sequence 804, Application US/10291265  
; Publication No. US2003023054A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang et al  
; TITLE OF INVENTION: No. US2003023054A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-017 (785)  
; CURRENT APPLICATION NUMBER: US/10/291,265  
; CURRENT FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/617,746  
; PRIOR FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/631,451  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 09/633,870  
; PRIOR FILING DATE: 2000-09-15  
; NUMBER OF SEQ ID NOS: 944  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 804  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-265-804

Query Match 90.2%; Score 37; DB 15; Length 384;  
Best Local Similarity 87.5%; Pred. No. 84;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 QQYSSSPT 8  
Db 259 QQYSSSPT 266

RESULT 2  
US-10-291-265-805  
; Sequence 805, Application US/10291265  
; Publication No. US20030232054A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang et al  
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-017 (785)  
; CURRENT APPLICATION NUMBER: US/10/291,265  
; CURRENT FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/617,746  
; PRIOR FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/631,451  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 09/633,870  
; PRIOR FILING DATE: 2000-09-15  
; NUMBER OF SEQ ID NOS: 944  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 805  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-265-805

Query Match 90.2%; Score 37; DB 15; Length 384;  
Best Local Similarity 87.5%; Pred. No. 84;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 QOYSSSPT 8  
||| ||||  
Db 259 QOYSSSPT 266

RESULT 3  
US-10-291-265-806  
; Sequence 806, Application US/10291265  
; Publication No. US20030232054A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang et al  
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-017 (785)  
; CURRENT APPLICATION NUMBER: US/10/291,265  
; CURRENT FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/617,746  
; PRIOR FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/631,451  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 09/633,870  
; PRIOR FILING DATE: 2000-09-15  
; NUMBER OF SEQ ID NOS: 944  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 806  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-265-806

Query Match 90.2%; Score 37; DB 15; Length 384;  
Best Local Similarity 87.5%; Pred. No. 84;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 QOYSSSPT 8  
||| ||||  
Db 259 QOYSSSPT 266

RESULT 4  
US-10-291-265-807  
; Sequence 807, Application US/10291265  
; Publication No. US20030232054A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang et al  
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-017 (785)  
; CURRENT APPLICATION NUMBER: US/10/291,265  
; CURRENT FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/617,746  
; PRIOR FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/631,451  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 09/633,870  
; PRIOR FILING DATE: 2000-09-15  
; NUMBER OF SEQ ID NOS: 944  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 807  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-265-807

Query Match 90.2%; Score 37; DB 15; Length 384;  
Best Local Similarity 87.5%; Pred. No. 84;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 QOYSSSPT 8  
||| ||||  
Db 259 QOYSSSPT 266

RESULT 5  
US-10-714-353-22  
; Sequence 22, Application US/10714353  
; Publication No. US20040170626A1  
; GENERAL INFORMATION:  
; APPLICANT: Schuurman, Janine  
; APPLICANT: Havenith, Karin  
; APPLICANT: Parren, Paul  
; APPLICANT: van de Winkel, Jan  
; APPLICANT: Williams, Denise Leah  
; APPLICANT: Jorgen Petersen  
; APPLICANT: Ole Baadsgaard  
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES AGAINST CD25  
; FILE REFERENCE: GMI-059  
; CURRENT APPLICATION NUMBER: US/10/714,353  
; CURRENT FILING DATE: 2003-11-14  
; PRIOR APPLICATION NUMBER: 60/426690  
; PRIOR FILING DATE: 2002-11-15  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-714-353-22

Query Match 87.8%; Score 36; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QOYSSSP 7  
||| ||||  
Db 1 QOYSSSP 7

RESULT 6

US-10-714-353-34  
; Sequence 34, Application US/10714353  
; Publication No. US20040170626A1  
; GENERAL INFORMATION:  
; APPLICANT: Schuurman, Janine  
; APPLICANT: Havenith, Karin  
; APPLICANT: Parren, Paul  
; APPLICANT: van de Winkel, Jan  
; APPLICANT: Williams, Denise Leah  
; APPLICANT: Jorgen Petersen  
; APPLICANT: Ole Baadsgaard  
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES AGAINST CD25  
; FILE REFERENCE: GMI-059  
; CURRENT APPLICATION NUMBER: US/10/714,353  
; PRIOR FILING DATE: 2003-11-14  
; PRIOR APPLICATION NUMBER: 60/426690  
; PRIOR FILING DATE: 2002-11-15  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-714-353-34

Query Match 87.8%; Score 36; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYSSSP 7  
Db 1 QYSSSP 7

RESULT 7  
US-10-714-353-40  
; Sequence 40, Application US/10714353  
; Publication No. US20040170626A1  
; GENERAL INFORMATION:  
; APPLICANT: Schuurman, Janine  
; APPLICANT: Havenith, Karin  
; APPLICANT: Parren, Paul  
; APPLICANT: van de Winkel, Jan  
; APPLICANT: Williams, Denise Leah  
; APPLICANT: Jorgen Petersen  
; APPLICANT: Ole Baadsgaard  
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES AGAINST CD25  
; FILE REFERENCE: GMI-059  
; CURRENT APPLICATION NUMBER: US/10/714,353  
; PRIOR FILING DATE: 2003-11-14  
; PRIOR APPLICATION NUMBER: 60/426690  
; PRIOR FILING DATE: 2002-11-15  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-714-353-40

Query Match 87.8%; Score 36; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYSSSP 7  
Db 1 QYSSSP 7

RESULT 8  
US-10-783-311-234  
; Sequence 234, Application US/10783311  
; Publication No. US20050009136A1

; GENERAL INFORMATION:  
; APPLICANT: Nixon, Andrew  
; APPLICANT: Hogan, Shannon  
; TITLE OF INVENTION: PAPP-A LIGANDS  
; FILE REFERENCE: 10280-059001  
; CURRENT APPLICATION NUMBER: US/10/783,311  
; CURRENT FILING DATE: 2004-02-19  
; PRIOR APPLICATION NUMBER: US 60/448,515  
; PRIOR FILING DATE: 2003-02-19  
; NUMBER OF SEQ ID NOS: 394  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 234  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Light Chain amino acid sequence  
US-10-783-311-234

Query Match 87.8%; Score 36; DB 17; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYSSSP 7  
Db 1 QYSSSP 7

RESULT 9  
US-10-723-434-16  
; Sequence 16, Application US/10723434  
; Publication No. US20040133357A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhong, Pingyu  
; APPLICANT: Luo, Peizhi  
; APPLICANT: Wang, Kevin C.  
; APPLICANT: Hsieh, Mark  
; APPLICANT: Li, Yan  
; TITLE OF INVENTION: HUMANIZED ANTIBODIES AGAINST VASCULAR ENDOTHELIAL GROWTH FACTOR  
; FILE REFERENCE: 26050-709.501  
; CURRENT APPLICATION NUMBER: US/10/723,434  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: US 60/284,407  
; PRIOR FILING DATE: 2001-04-17  
; PRIOR APPLICATION NUMBER: US 10/125,687  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: US 10/153,176  
; PRIOR FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: US 10/443,134  
; PRIOR FILING DATE: 2003-05-20  
; NUMBER OF SEQ ID NOS: 156  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: VL  
US-10-723-434-16

Query Match 87.8%; Score 36; DB 16; Length 107;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYSSSP 7  
Db 89 QYSSSP 95

RESULT 10  
US-10-783-311-230  
; Sequence 230, Application US/10783311  
; Publication No. US20050009136A1

```

; GENERAL INFORMATION:
; APPLICANT: Nixon, Andrew
; APPLICANT: Hogan, Shannon
; TITLE OF INVENTION: PAPP-A LIGANDS
; FILE REFERENCE: 10280-059001
; CURRENT APPLICATION NUMBER: US/10/783,311
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: US 60/448,515
; PRIOR FILING DATE: 2003-02-19
; NUMBER OF SEQ ID NOS: 394
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 230
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Light Chain amino acid sequence
US-10-783-311-230

```

```

Query Match      87.8%; Score 36; DB 17; Length 116;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 QOYSSSP 7
   |||||
Db 90 QOYSSSP 96

```

```

RESULT 11
US-10-714-353-4
; Sequence 4, Application US/10714353
; Publication No. US20040170626A1
; GENERAL INFORMATION:
; APPLICANT: Schuurman, Janine
; APPLICANT: Havenith, Karin
; APPLICANT: Parren, Paul
; APPLICANT: van de Winkel, Jan
; APPLICANT: Williams, Denise Leah
; APPLICANT: Jorgen Petersen
; APPLICANT: Ole Baadsgaard
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES AGAINST CD25
; FILE REFERENCE: GMI-059
; CURRENT APPLICATION NUMBER: US/10/714,353
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 60/426690
; PRIOR FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-714-353-4

```

```

Query Match      87.8%; Score 36; DB 16; Length 120;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 QOYSSSP 7
   |||||
Db 90 QOYSSSP 96

```

```

RESULT 12
US-10-714-353-12
; Sequence 12, Application US/10714353
; Publication No. US20040170626A1
; GENERAL INFORMATION:
; APPLICANT: Schuurman, Janine
; APPLICANT: Havenith, Karin
; APPLICANT: Parren, Paul
; APPLICANT: van de Winkel, Jan
; APPLICANT: Williams, Denise Leah

```

```

; APPLICANT: Jorgen Petersen
; APPLICANT: Ole Baadsgaard
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES AGAINST CD25
; FILE REFERENCE: GMI-059
; CURRENT APPLICATION NUMBER: US/10/714,353
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 60/426690
; PRIOR FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-714-353-12

```

```

Query Match      87.8%; Score 36; DB 16; Length 120;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 QOYSSSP 7
   |||||
Db 90 QOYSSSP 96

```

```

RESULT 13
US-10-365-761B-19
; Sequence 19, Application US/10365761B
; Publication No. US20040023300A1
; GENERAL INFORMATION:
; APPLICANT: Grabe, Niels
; APPLICANT: Konig, Matthias
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF A FUNCTIONAL
; TITLE OF INVENTION: PROTEIN SEQUENCE AND AN APPARATUS THEREFOR
; FILE REFERENCE: AEOHW7.001AUS
; CURRENT APPLICATION NUMBER: US/10/365,761B
; CURRENT FILING DATE: 2003-04-12
; PRIOR APPLICATION NUMBER: EP 02003469.0
; PRIOR FILING DATE: 2002-02-14
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthesized proven phosphorylation site
US-10-365-761B-19

```

```

Query Match      80.5%; Score 33; DB 15; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.6e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 QOYSSSPT 8
   |||||
Db 2 QRYSSDPT 9

```

```

RESULT 14
US-10-862-195-1950
; Sequence 1950, Application US/10862195
; Publication No. US20050164324A1
; GENERAL INFORMATION:
; APPLICANT: GYGI, STEVEN P.
; TITLE OF INVENTION: SYSTEMS, METHODS AND KITS FOR CHARACTERIZING PHOSPHOPROTEOMES
; FILE REFERENCE: 58890(70207)
; CURRENT APPLICATION NUMBER: US/10/862,195
; CURRENT FILING DATE: 2004-06-04
; PRIOR APPLICATION NUMBER: 60/476,010
; PRIOR FILING DATE: 2003-06-04
; NUMBER OF SEQ ID NOS: 2245
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1950

```

; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-10-862-195-1950

Query Match 80.5%; Score 33; DB 18; Length 11;  
Best Local Similarity 85.7%; Pred. No. 13;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYSSSP 7  
Db 3 QYSSAP 9

RESULT 15  
US-10-437-963-185574  
; Sequence 185574, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 185574  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(68)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_82457C.1.pep  
US-10-437-963-185574

Query Match 80.5%; Score 33; DB 16; Length 68;  
Best Local Similarity 85.7%; Pred. No. 84;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYSSSPT 8  
Db 24 EYSSSPT 30

Search completed: August 9, 2005, 14:26:50  
Job time : 87.4615 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:06:08 ; Search time 17.8462 Seconds  
(without alignments)  
43.132 Million cell updates/sec

Title: US-10-089-452-29

Perfect score: 41

Sequence: 1 QQYSSSPT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	90.2	91	2 S37520	Ig kappa chain V r
2	37	90.2	108	2 H44151	Ig kappa chain V r
3	36	87.8	90	2 I38601	Ig kappa chain V r
4	33	80.5	87	2 T18182	hypothetical prote
5	33	80.5	109	1 K3HUT1	Ig kappa chain V-I
6	33	80.5	109	2 F44151	Ig kappa chain V r
7	33	80.5	128	1 K3HU41	Ig kappa chain pre
8	33	80.5	604	1 TVYU4H	protein-tyrosine k
9	33	80.5	698	1 TVFVLV	protein-tyrosine k
10	33	80.5	1210	1 GQHUE	epidermal growth f
11	33	80.5	1210	2 A53183	epidermal growth f
12	33	80.5	1223	1 TVCHLV	epidermal growth f
13	33	80.5	1733	1 RNBY2L	DNA-directed RNA p
14	32	78.0	54	2 A25521	Ig kappa chain V r
15	32	78.0	91	2 S67940	Ig kappa chain V r
16	32	78.0	92	2 S37513	Ig kappa chain V r
17	32	78.0	92	2 S37510	Ig kappa chain V r
18	32	78.0	92	2 S37518	Ig kappa chain V r
19	32	78.0	92	2 S37517	Ig kappa chain V r
20	32	78.0	92	2 S37524	Ig kappa chain V r
21	32	78.0	92	2 S37519	Ig kappa chain V r
22	32	78.0	93	2 S37526	Ig kappa chain V r
23	32	78.0	93	2 S37501	Ig kappa chain V r
24	32	78.0	93	2 S37528	Ig kappa chain V r
25	32	78.0	96	2 A30601	Ig kappa chain V-I
26	32	78.0	107	2 S36275	Ig lambda chain V
27	32	78.0	108	1 K3HUB6	Ig kappa chain V-I
28	32	78.0	108	2 B30608	Ig kappa chain V-I
29	32	78.0	108	2 C30608	Ig kappa chain V-I

30	32	78.0	109	1 K3HUGO	Ig kappa chain V-I
31	32	78.0	109	1 K3HUSI	Ig kappa chain V-I
32	32	78.0	109	2 S47181	Ig kappa chain - h
33	32	78.0	109	2 H30601	Ig kappa chain V-I
34	32	78.0	109	2 G30607	Ig kappa chain V-I
35	32	78.0	109	2 A30608	Ig kappa chain V-I
36	32	78.0	109	2 G30601	Ig kappa chain V-I
37	32	78.0	109	2 C30601	Ig kappa chain V-I
38	32	78.0	109	2 B30601	Ig kappa chain V-I
39	32	78.0	109	2 PH0963	Ig kappa chain V r
40	32	78.0	109	2 D30601	Ig kappa chain V-I
41	32	78.0	111	2 S40359	Ig kappa chain V-J
42	32	78.0	114	2 S46375	Ig kappa chain V-J
43	32	78.0	116	2 B27594	Ig kappa chain pre
44	32	78.0	116	2 B25521	Ig kappa chain pre
45	32	78.0	121	2 S40327	Ig kappa chain - h

#### ALIGNMENTS

##### RESULT 1

S37520

Ig kappa chain V region (V-kappa 3) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999

C;Accession: S37520

R;Klein, U.; Kueppers, R.; Rajewsky, K.

submitted to the EMBL Data Library, September 1993

A;Description: Human IGM(+) Igd(+) cells, the major B cell subset in the peripheral blood

A;Reference number: S37501

A;Accession: S37520

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-91 <KLE>

A;Cross-references: EMBL:226612; NID:g405682; PIDN:CAA81365.1; PID:g405683

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotrimer; immunoglobulin

Query Match 90.2%; Score 37; DB 2; Length 91;

Best Local Similarity 87.5%; Pred. No. 1.1;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QQYSSSPT 8

Db 73 QQYSSSPT 80

##### RESULT 2

H44151

Ig kappa chain V region (JW-15) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 09-Jul-2004

C;Accession: H44151

R;Zebede, S.L.; Barbas III, C.F.; Hom, Y.L.; Caethien, R.H.; Graff, R.; DeGraw, J.; Pya

Proc. Natl. Acad. Sci. U.S.A. 89, 3175-3179, 1992

A;Title: Human combinatorial antibody libraries to hepatitis B surface antigen.

A;Reference number: A44151; MUID:92228746; PMID:1373487

A;Accession: H44151

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-108 <ZEB>

A;Cross-references: UNIPROT:Q9UL78

A;Note: nucleotide translation not given

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotrimer; immunoglobulin

F;15-90/Domain: immunoglobulin homology <IMM>

Query Match 90.2%; Score 37; DB 2; Length 108;

Best Local Similarity 87.5%; Pred. No. 1.3;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QQYSSSPT 8

```
Db          89 QOYSSSPT 96
          ||| |||
RESULT 3
Ig kappa chain V region (5F8) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 23-Jul-1999
C:Accession: I38601
R:Goshorn, S.C.; Retzel, E.; Jemerson, R.
J. Biol. Chem. 266, 2134-2142, 1991
A:Title: Common structural features among monoclonal antibodies binding the same antigen
A:Reference number: A38601; MUID:91115823; PMID:1703527
A:Accession: I38601
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-90 <GOS>
A:Cross-references: GB:M57986; NID:g196418; PIDN:AAA63367.1; PID:g196419
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match      87.8%; Score 36; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYSSSP 7
    |||||
Db 71 QOYSSSP 77

RESULT 4
T18182
Hypothetical protein a680R - Chlorella virus PBCV-1
C:Species: Chlorella virus PBCV-1
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18182
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: T18182
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-87 <GRA>
A:Cross-references: UNIPROT:O41162; EMBL:U42580; NID:g4028896; PIDN:AAC97053.1
A:Experimental source: specific host Chlorella strain NC64A
C:Genetics:
A:Note: a680R

Query Match      80.5%; Score 33; DB 2; Length 87;
Best Local Similarity 75.0%; Pred. No. 7.3;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 1 QOYSSSPT 8
    |||||
Db 77 QOYSSKPT 84

RESULT 5
K3HUTI
Ig kappa chain V-III region (Ti) - human
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 09-Jul-2004
C:Accession: A01895
R:Stuter, L.; Barnikol, H.U.; Watanabe, S.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 353, 189-208, 1972
A:Title: Die Primärstruktur einer monoklonalen Immunglobulin-L-Kette vom kappa-Typ, Sub
A:Reference number: A91651; MUID:72188439; PMID:5027703
A:Accession: A01895
A:Molecule type: protein
A:Residues: 1-109 <SUT>
A:Cross-references: UNIPROT:P01622
A:Note: the sequence of the C region, which has the Inv (3) marker, is also given
```

```
C:Comment: This is a Bence Jones protein.
C:Genetics:
A:Gene: GDB:IGKV3
A:Cross-references: GDB:I36266
A:Map position: 2p12-2p11
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into larger superfamilies: immunoglobulin V region; immunoglobulin homology
C:Superfamily: heterotetramer
C:Keywords: heterotetramer
F:16-91/Domain: immunoglobulin homology <IMM>
F:23-89/Disulfide bonds: #status predicted

Query Match      80.5%; Score 33; DB 1; Length 109;
Best Local Similarity 75.0%; Pred. No. 9.2;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYSSSPT 8
    |||||
Db 90 QOYSSSPS 97

RESULT 6
F44151
Ig kappa chain V region (JM-01) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 21-Jan-2000
C:Accession: F44151
R:Zebede, S.L.; Barbas III, C.F.; Hom, Y.L.; Caothien, R.H.; Graff, R.; DeGraw, J.; Pyat
Proc. Natl. Acad. Sci. U.S.A. 89, 3175-3179, 1992
A:Title: Human combinatorial antibody libraries to hepatitis B surface antigen.
A:Reference number: A44151; MUID:92228746; PMID:1373487
A:Accession: F44151
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-109 <ZEB>
A:Cross-references: GB:M88314; NID:g183962; PIDN:AAA35972.1; PID:g183963
A:Note: nucleotide translation not given
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-90/Domain: immunoglobulin homology <IMM>

Query Match      80.5%; Score 33; DB 2; Length 109;
Best Local Similarity 75.0%; Pred. No. 9.2;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYSSSPT 8
    |||||
Db 89 QOYSSSPS 96

RESULT 7
K3HU41
Ig kappa chain precursor V-III region (IARC/BL41) - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A01899
R:Klobeck, H.G.; Meindl, A.; Combiato, G.; Solomon, A.; Zachau, H.G.
Nucleic Acids Res. 13, 6499-6513, 1985
A:Title: Human immunoglobulin kappa light chain genes of subgroups II and III.
A:Reference number: A93588; MUID:86041852; PMID:2997711
A:Accession: A01899
A:Molecule type: DNA
A:Residues: 1-128 <KLO>
A:Cross-references: UNIPROT:P06311
A:Note: the sequence was determined from the differentiated gene
C:Genetics:
A:Gene: GDB:IGKV3
A:Cross-references: GDB:I36266
A:Map position: 2p12-2p11
A:Introns: 17/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into larger superfamilies: immunoglobulin V region; immunoglobulin homology
```



C;Keywords: heterotrimer; immunoglobulin  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-128/Product: Ig kappa chain V-III region (IARC/BL41) #status predicted <MAT>  
F;21-43/Region: framework 1  
F;36-110/Domain: immunoglobulin homology <IMW>  
F;44-54/Region: complementarity-determining 1  
F;55-69/Region: framework 2  
F;70-76/Region: complementarity-determining 2  
F;77-108/Region: framework 3  
F;109-117/Region: complementarity-determining 3  
F;118-128/Region: framework 4  
F;43-108/Disulfide bonds: #status predicted

Query Match 80.5%; Score 33; DB 1; Length 128;  
Best Local Similarity 85.7%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYSSSP 7  
|:|:|:|  
Db 109 QYSTSP 115

RESULT 8  
TVVUH  
protein-tyrosine kinase (EC 2.7.1.112) erbB - avian erythroblastosis virus (strain H)  
C;Species: avian erythroblastosis virus  
C;Date: 18-Apr-1984 #sequence\_revision 18-Apr-1984 #text\_change 09-Jul-2004  
C;Accession: A00644; A38022  
R;Yamamoto, T.; Nishida, T.; Miyajima, N.; Kawai, S.; Ooi, T.; Toyoshima, K.  
Cell 35, 71-78, 1983  
A;Title: The erbB gene of avian erythroblastosis virus is a member of the src gene family  
A;Reference number: A00644; MUID:84026539; PMID:6313229  
A;Accession: A00644

A;Molecule type: DNA  
A;Residues: 1-604 <YAM>  
A;Cross-references: UNIPROT:P00535; GB:K01216; NID:g209676; PIDN:AAA42400.1; PID:g209678  
R;Debuire, B.; Henry, C.; Benaisse, M.; Biserte, G.; Claverie, J.M.; Saule, S.; Martin, Science 224, 1456-1459, 1984  
A;Title: Sequencing the erbA gene of avian erythroblastosis virus reveals a new type of  
A;Reference number: A38022; MUID:8423957; PMID:6328658  
A;Accession: A38022

A;Molecule type: DNA  
A;Residues: 1-28,'W',30-139,'F',141-145,'V',147-152 <DEB>  
A;Cross-references: GB:K02006  
C;Genetics:  
A;Gene: erbB  
C;Superfamily: epidermal growth factor receptor; protein kinase homology  
C;Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific P  
F;130-395/Domain: protein kinase homology <KIN>  
F;138-146/Region: protein kinase ATP-binding motif  
F;165/Active site: Lys #status predicted

Query Match 80.5%; Score 33; DB 1; Length 604;  
Best Local Similarity 75.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYSSSP 8  
|:|:|:|  
Db 486 QYSSDPT 493

RESULT 9  
TVFLV  
protein-tyrosine kinase (EC 2.7.1.112) erbB - avian leukosis virus  
N;Contains: amino end of gag protein; env protein fragment; protein-tyrosine kinase  
C;Species: avian leukosis virus, ALV  
C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C;Accession: B00643; A00643  
R;Nilsson, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines, M Cell 41, 719-726, 1985  
A;Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and pro  
A;Reference number: A00643; MUID:85228222; PMID:2988784  
A;Accession: B00643

A;Molecule type: mRNA  
A;Residues: 1-698 <NIL>  
A;Cross-references: UNIPROT:P00534; GB:M10066; GB:M13881; NID:g211749; PIDN:AAA48763.1; I  
A;Note: in Genbank entry CHKRRBPF, release 109.0, the source is designated as Gallus gall  
C;Comment: This protein is synthesized as a gag-env-erbB protein.  
C;Genetics:  
A;Gene: gag-env-erbB  
C;Superfamily: epidermal growth factor receptor; protein kinase homology  
C;Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific p  
F;1-6/Product: gag protein (fragment) #status predicted <GAG>  
F;7-59/Product: env protein (fragment) #status predicted <ENV>  
F;60-698/Product: protein-tyrosine kinase erbB #status predicted <ERB>  
F;194-459/Domain: protein kinase homology <KIN>  
F;202-210/Region: protein kinase ATP-binding motif  
F;229/Active site: Lys #status predicted

Query Match 80.5%; Score 33; DB 1; Length 698;  
Best Local Similarity 75.0%; Pred. No. 63;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYSSSP 8  
|:|:|:|  
Db 550 QYSSDPT 557

RESULT 10  
GQHUE  
epidermal growth factor receptor precursor - human  
N;Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB  
C;Species: Homo sapiens (man)  
C;Date: 15-Nov-1984 #sequence\_revision 27-Nov-1985 #text\_change 09-Jul-2004  
C;Accession: A00641; A25772; S30024; A38672; A00642; A43615; A23062; A05281; A60143; A33  
R;Ullrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Ye  
rg, P.H.  
Nature 309, 418-425, 1984  
A;Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression of  
A;Reference number: A00641; MUID:84219729; PMID:6328312  
A;Accession: A00641

A;Molecule type: mRNA  
A;Residues: 1-1210 <UL>  
A;Cross-references: UNIPROT:P00533; EMBL:X00588; NID:g31113; PIDN:CAA52540.1; PID:g75792  
A;Note: the authors translated the codon AAG for residue 540 as Asn  
R;Ishii, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I. Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985  
A;Title: Characterization and sequence of the promoter region of the human epidermal gro  
A;Reference number: A25772; MUID:85270438; PMID:2991899  
A;Accession: A25772  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-29 <ISH>  
A;Cross-references: GB:M11234; NID:gi181981; PIDN:AAAS2370.1; PID:g553272  
R;Haley, J.; Whittle, N.; Bennett, P.; Kinchington, D.; Ullrich, A.; Waterfield, M. Oncogene Res. 1, 375-396, 1987  
A;Title: The human EGF receptor gene: structure of the 110 kb locus and identification o  
A;Reference number: S30024; MUID:88217333; PMID:3329716  
A;Accession: S30024

A;Molecule type: DNA  
A;Residues: 1-29 <HA2>  
A;Cross-references: EMBL:X06370; NID:g31118; PIDN:CAA29668.1; PID:g31119  
R;Haley, J.D.; Waterfield, M.D. J. Biol. Chem. 266, 1746-1753, 1991  
A;Title: Contributory effects of de Novo transcription and premature transcript terminat  
A;Reference number: A38672; MUID:91107677; PMID:1988448  
A;Accession: A38672

A;Molecule type: DNA  
A;Residues: 1-29 <HAL>  
A;Cross-references: GB:M38425; NID:gi81977; PIDN:AAA63171.1; PID:g553271  
A;Experimental source: carcinoma cell line A431-7  
R;Xu, Y.; Ishii, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.; Mer Nature 309, 806-810, 1984  
A;Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs  
A;Reference number: A00642; MUID:84245835; PMID:6330563  
A;Accession: A00642

A:Molecule type: mRNA  
A:Residues: 'RCAMRA',150-187,'KSVIQAV',195,'M',197,'A',199-222,'S',224-304,'RA',307-321  
,798-799,'TD',802-811,'R',813-942 <XUY>  
A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF receptor  
R:Lin, C.R.; Chen, W.S.; Kruijer, W.; Stolarsky, L.S.; Weber, W.; Evans, R.M.; Verma, I.M.  
Science 224, 843-848, 1984  
A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplification  
A:Reference number: A43615; MUID:84196372; PMID:6326261  
A:Accession: A43615  
A:Molecule type: mRNA  
A:Residues: 713-964 <LIN>  
A:Experimental source: epidermoid carcinoma cell line A431  
R:Simmen, F.A.; Gope, M.L.; Schulz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.  
Biochem. Biophys. Res. Commun. 124, 125-132, 1984  
A:Reference number: A23062; MUID:85046483; PMID:6093780  
A:Accession: A23062  
A:Molecule type: mRNA  
A:Residues: 1028-1210 <SIM>  
R:Weber, W.; Gill, G.N.; Speiss, J.  
Science 224, 294-297, 1984  
A:Reference number: A05281; MUID:84172183; PMID:6324343  
A:Accession: A05281  
A:Molecule type: protein  
A:Residues: 25-30,'S',32-51;454-467 <WEB>  
R:Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.  
J. Biol. Chem. 260, 5205-5208, 1985  
A:Title: Identification of residues in the nucleotide binding site of the epidermal growth  
A:Reference number: A60143; MUID:85182650; PMID:2985580  
A:Accession: A60143  
A:Molecule type: protein  
A:Residues: 740-744,'X',746-747 <RUS>  
R:Mróczkowski, B.; Mosig, G.; Cohen, S.  
Nature 309, 270-273, 1984  
A:Title: ATP-stimulated interaction between epidermal growth factor receptor and supercoiled  
A:Reference number: A38023; MUID:84191554; PMID:6325948  
A:Contents: annotation; receptor activity  
A:Note: The EGF receptor (and other tyrosine kinases) can nick double-stranded DNA  
R:Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.  
Cell 59, 33-43, 1989  
A:Title: Functional independence of the epidermal growth factor receptor from a domain  
A:Reference number: A33331; MUID:90003233; PMID:2790960  
A:Contents: annotation; internalization signal  
C:Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor  
C:Genetics:  
A:Gene: GDB:EGFR  
A:Cross-references: GDB:120610; OMIM:131550  
A:Map position: Tp12.3-7p12.1  
C:Superfamily: epidermal growth factor receptor; protein kinase homology  
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-1210/Product: EGF receptor #status predicted <MAT>  
F:25-645/Domain: extracellular #status predicted <EXT>  
F:73-300/Domain: EGF receptor extracellular domain repeat <BE1>  
F:390-600/Domain: EGF receptor extracellular domain repeat <EE2>  
F:646-668/Domain: transmembrane #status predicted <TMW>  
F:669-1210/Domain: intracellular #status predicted <INT>  
F:710-975/Domain: protein kinase homology <KIN>  
F:718-726/Region: protein kinase ATP-binding motif  
F:939-1046/Region: coated-pit mediated internalization signal  
F:1047-1210/Region: inhibitory  
F:128,175,352,413,444,528,603/Binding site: carbohydrate (Asn) (covalent) #status predic  
F:745/Active site: Lys #status experimental

Query Match	80.5%	Score 33;	DB 1;	Length 1210;
Best Local Similarity	75.0%	Pred. NO. 1.1e+02;		
Matches	6;	Conservative	1;	Mismatches 1; Indels 0; Gaps 0;

Qy 1 QYSSSPT 8  
Db 1067 QRYSSDPT 1074

RESULT 11

A53183  
epidermal growth factor receptor precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 09-Jul-2004  
C:Accession: A53183; A43818; S24942; A28941; S45325; I45643  
R:Luettker, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.; I  
Genes Dev. 8, 399-413, 1994  
A:Title: The mouse wavy-2 phenotype results from a point mutation in the EGF receptor ty  
A:Reference number: A53183; MUID:94170986; PMID:8125255  
A:Accession: A53183  
A:Molecule type: mRNA  
A:Residues: 1-1210 <LUE>  
A:Cross-references: UNIPROT:Q01279; GB:U03425  
R:Avivi, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.  
Oncogene 6, 673-676, 1991  
A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding site  
A:Reference number: A43818; MUID:91232866; PMID:2030916  
A:Accession: A43818  
A:Molecule type: mRNA  
A:Residues: 1-714 <AVI>  
A:Cross-references: GB:X59698  
R:Eisinger, D.P.; Serrero, G.  
submitted to the EMBL Data Library, June 1992  
A:Reference number: S24942  
A:Accession: S24942  
A:Molecule type: mRNA  
A:Residues: 969-971, 'K', 973-1115, 'D' <EIS>  
A:Cross-references: EMBL:Z12608  
R:Heisermann, G.J.; Gill, G.N.  
J. Biol. Chem. 263, 13152-13158, 1988  
A:Title: Epidermal growth factor receptor threonine and serine residues phosphorylated ir  
A:Reference number: A28941; MUID:88330814; PMID:3138233  
A:Accession: A28941  
A:Molecule type: protein  
A:Residues: 689-694, 'X', 696-704, 'L', 706-707; 989-992, 'XX', 995-996, 'X', 998-1000; 1002-1009, '  
R:Habbs, M.L.; Dunn, A.R.; Alexander, W.S.  
submitted to the EMBL Data Library, April 1994  
A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor  
A:Reference number: S45325  
A:Accession: S45325  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-971, 'K', 973-1210 <VER>  
A:Cross-references: EMBL:X78987; NID:G488830; PIDN:CAA55587.1; PID:G488831  
R:Paria, B.C.; Das, S.K.; Andrews, G.K.; Dev, S.K.  
Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993  
A:Title: Expression of the epidermal growth factor receptor gene is regulated in mouse b  
A:Reference number: I49643; MUID:93126380; PMID:7678348  
A:Accession: I49643  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 12-20,22-132 <RES>  
A:Cross-references: GB:L06864; NID:G193001; PIDN:AAA53029.1; PID:G567201  
C:Genetics:  
A:Gene: EGFR  
C:Superfamily: epidermal growth factor receptor; protein kinase homology  
C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phosphop  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:648-670/Domain: transmembrane #status predicted <TM>  
F:712-977/Domain: protein kinase homology <KIN>  
F:720-728/Region: protein kinase ATP-binding motif  
F:680,695/Binding site: phosphate (Thr) (covalent) #status experimental  
F:697,1070,1071/Binding site: phosphate (Ser) (covalent) #status experimental  
F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental  
F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental  
F:1197/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 80.5%; Score 33; DB 2; Length 1210;  
Best Local Similarity 75.0%; Pred No. 1,1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QQQSSSPT 8  
|||||

Db 1067 QRYSSDPT 1074

RESULT 12

TVCHLV

epidermal growth factor receptor precursor - chicken

N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB

C:Species: Gallus gallus (Chicken)

C>Date: 28-Feb-1986 #sequence\_revision 05-May-1995 #text\_change 09-Jul-2004

C:Accession: A27720; A00643

R:Fax, I.; Johnson, A.; Hawk, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Vennstr

Mol. Cell. Biol. 8, 1970-1978, 1988

A:Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in mou

A:Reference number: A27720; MUID:88261272; PMID:3260329

A:Accession: A27720

A:Molecule type: mRNA

A:Residues: 1-1223 <LAX>

A:Cross-references: UNIPROT:P00534; GB:M20386

R:Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines, M

Cell 41, 719-726, 1985

A:Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and pro

A:Reference number: A00643; MUID:85228222; PMID:2988784

A:Accession: A00643

A:Molecule type: mRNA

A:Residues: 585-1223 <NIL>

A:Cross-references: GB:M10066

C:Genetics:

A:Gene: erbB

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor

pecific protein kinase

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-1223/Product: epidermal growth factor receptor #status predicted <MAT>

F:31-654/Domain: extracellular #status predicted <EXT>

F:81-307/Domain: EGF receptor extracellular domain repeat <EE1>

F:397-610/Domain: EGF receptor extracellular domain repeat <EE2>

F:655-677/Domain: transmembrane #status predicted <TM>

F:678-1223/Domain: intracellular #status predicted <INT>

F:719-984/Domain: protein kinase homology <KIN>

F:727-735/Region: protein kinase ATP-binding motif

F:136,202,280,361,370,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent) #e

F:192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted

F:687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

F:754/Active site: Lys #status predicted

F:1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Query Match 80.5%; Score 33; DB 1; Length 1223;

Best Local Similarity 75.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQYSSSPT 8

Db 1075 QRYSSDPT 1082

RESULT 13

RNBV2L

DNA-directed RNA polymerase (EC 2.7.7.6) II 215K chain - yeast (Saccharomyces cerevisiae

N:Alternate names: protein D2150; protein YDL140c; RNA polymerase II largest chain

C:Species: Saccharomyces cerevisiae

C>Date: 04-Dec-1986 #sequence\_revision 23-Aug-1996 #text\_change 09-Jul-2004

C:Accession: S67686; A00692; S64647; S20985

R:Saluz, H.P.; Woelfl, S.; Hansmann, V.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67677

A:Accession: S67686

A:Molecule type: DNA

A:Residues: 1-1733 <SAL>

A:Cross-references: UNIPROT:P04050; EMBL:Z74188; NID:g1431216; PIDN:CAA98713.1; PID:g143

A:Experimental source: strain S288C

R:Allison, L.A.; Moyle, M.; Shales, M.; Ingles, C.J.

Cell 42, 599-610, 1985

A:Title: Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA

A:Reference number: A90870; MUID:85282617; PMID:3896517

A:Accession: A00692

A:Molecule type: DNA

A:Residues: 1-1513, 'V', 1515-1523, 'A', 1525-1555, 1563-1607, 'M', 1609-1733 <ALL>

A:Cross-references: EMBL:X03128; NID:g4397; PIDN:CAA26904.1; PID:g4398

R:Cronan Jr., J.E.; Wallace, J.C.

FEMS Microbiol. Lett. 130, 221-230, 1995

A:Title: The gene encoding the biotin-apoptrotein ligase of Saccharomyces cerevisiae.

A:Reference number: S64646; MUID:95377607; PMID:7649444

A:Accession: S64647

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1669-1733 <CRO>

A:Cross-references: EMBL:U27182; NID:g886080; PIDN:AAC49058.1; PID:g886082

C:Genetics:

A:Gene: SGD:RPO21; MIPS:S0002299; MIPS:YDL140C

A:Cross-references: SGD:S0002299; MIPS:YDL140C

A:Map position: 4L

C:Superfamily: human DNA-directed RNA polymerase II largest chain

C:Keywords: DNA binding; nucleotidyltransferase; tandem repeat; transcription; zinc finger

F:1567-1733/Region: 7-residue repeats

Query Match 80.5%; Score 33; DB 1; Length 1733;

Best Local Similarity 85.7%; Pred. No. 1.6e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQYSSSP 7

Db 4 QQYSSAP 10

RESULT 14

A25521

IG kappa chain V region (321) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 21-May-1988 #sequence\_revision 21-May-1990 #text\_change 09-May-1997

C:Accession: A25521

R:Chen, P.P.; Albrandt, K.; Orida, N.K.; Radoux, V.; Chen, E.Y.; Schrantz, R.; Liu, F.T.;

Proc. Natl. Acad. Sci. U.S.A. 83, 8318-8322, 1986

A:Title: Genetic basis for the cross-reactive idiotypes on the light chains of human IgM

A:Reference number: A94135; MUID:87041448; PMID:3095834

A:Accession: A25521

A:Molecule type: DNA

A:Residues: 1-54 <CHE>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 78.0%; Score 32; DB 2; Length 54;

Best Local Similarity 85.7%; Pred. No. 7.2;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQYSSSP 7

Db 48 QQYGSSP 54

RESULT 15

S67940

IG kappa chain V region, subgroup III (clone MH52) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 19-Mar-1997 #sequence\_revision 09-May-1997 #text\_change 20-Mar-1998

C:Accession: S67940

R:Hexham, J.M.; Furmaniak, J.; Pegg, C.; Burton, D.R.; Smith, B.R.

Autoimmunity 12, 135-141, 1992

A:Title: Cloning of a human autoimmune response: preparation and sequencing of a human ar

A:Reference number: S67940; MUID:92314301; PMID:1617110

A:Accession: S67940

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-91 <HEX>

A:Cross-references: EMBL:X73852

C:Superfamily: immunoglobulin V region; immunoglobulin homology

Query Match 78.0%; Score 32; DB 2; Length 91;  
Best Local Similarity 85.7%; Pred. No. 12;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QQYSSSP 7  
Db 71 QQYSSSP 77

Search completed: August 9, 2005, 14:20:40  
Job time : 17.8462 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2005, 14:05:28 ; Search time 85.8462 Seconds  
(without alignments)  
47.721 Million cell updates/sec

Title: US-10-089-452-29  
Perfect score: 41  
Sequence: 1 QQYSSSPT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	87.8	1230	2 Q9UVG1	Q9UVG1 tuber magna
2	35	85.4	230	2 Q6P0C7	Q6P0C7 brachydanio
3	34	82.9	1059	2 Q8BJ42	Q8BJ42 m mus muscu
4	34	82.9	1089	1 DLP2 RAT	P97837 rattus norv
5	33	80.5	87	2 Q41162	Q41162 paramesium
6	33	80.5	109	1 KV3D HUMAN	P01622 homo sapien
7	33	80.5	116	2 Q62954	Q62954 rattus norv
8	33	80.5	128	1 KV3K HUMAN	P06311 homo sapien
9	33	80.5	131	2 Q811C3	Q811C3 mus musculu
10	33	80.5	249	2 Q6NGH3	Q6NGH3 corynebacte
11	33	80.5	307	2 Q43416	Q43416 cenchrus ci
12	33	80.5	327	2 Q6ER51	Q6ER51 oryza sativ
13	33	80.5	339	2 Q8MV41	Q8MV41 aedes aegyp
14	33	80.5	387	2 Q75MF2	Q75MF2 homo sapien
15	33	80.5	509	2 Q6LHL4	Q6LHL4 photobacter
16	33	80.5	512	2 Q81FW9	Q81FW9 ciona inter
17	33	80.5	567	2 Q86714	Q86714 avian rous-
18	33	80.5	569	2 Q8RBG9	Q8RBG9 thermoanaer
19	33	80.5	579	2 Q86BM9	Q86BM9 drosophila
20	33	80.5	604	1 ERBB AVIER	P00534 avian eryth
21	33	80.5	634	1 ERBB ALV	P00534 avian leuko
22	33	80.5	665	2 Q94053	Q94053 candida alb
23	33	80.5	665	2 Q6SYI6	Q6SYI6 candida alb
24	33	80.5	729	2 Q86712	Q86712 avian rous-
25	33	80.5	826	2 Q81V15	Q81V15 homo sapien
26	33	80.5	845	2 Q7SED8	Q7SED8 nemospora
27	33	80.5	925	2 Q6CTI4	Q6CTI4 kluyveromyc
28	33	80.5	935	2 Q8MV44	Q8MV44 aedes aegyp
29	33	80.5	998	2 Q8MN47	Q8MN47 dictyostell
30	33	80.5	1002	2 Q6AVJ4	Q6AVJ4 oryza sativ
31	33	80.5	1209	2 Q8MIL8	Q8MIL8 sus scrofa

32	33	80.5	1209	2 Q9QX70	Q9QX70 rattus norv
33	33	80.5	1210	1 EGFR_HUMAN	P00533 homo sapien
34	33	80.5	1210	1 EGFR_MOUSE	Q01279 mus musculu
35	33	80.5	1210	2 Q9EP98	Q9EP98 mus musculu
36	33	80.5	1436	2 Q9LPY6	Q9LPY6 arabidopsis
37	33	80.5	1505	2 Q98TW1	Q98TW1 brachydanio
38	33	80.5	1715	2 Q6FQ89	Q6FQ89 candida gla
39	33	80.5	1727	2 Q68FD9	Q68FD9 mus musculu
40	33	80.5	1733	1 RPB1_YEAST	P04050 saccharomyc
41	32	78.0	108	1 KV3A HUMAN	P01619 homo sapien
42	32	78.0	109	1 KV3B HUMAN	P01620 homo sapien
43	32	78.0	109	1 KV3G_HUMAN	P04206 homo sapien
44	32	78.0	109	2 Q9UL78	Q9UL78 homo sapien
45	32	78.0	129	1 KV3W_HUMAN	P18136 homo sapien

ALIGNMENTS

RESULT 1  
Q9UVG1 PRELIMINARY; PRT; 1230 AA.  
AC Q9UVG1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Chitin synthase.  
GN Name=chs4;  
OS Tuber magnatum (white Piedmont truffle).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Pezizomycetes;  
OC Pezizales; Tuberaceae; Tuber.  
OX NCBI\_TaxID=42249;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Garnerio L., Lazzari B., Viotti A., Bonfante P.;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AUI31706; CAB41410.1; -;  
DR GO; GO:0016758; F:transferase activity, transferring hexosyl . . . ; IEA.  
DR InterPro; IPR004835; Chitin\_synth\_fng.  
DR InterPro; IPR001199; Cyt B5.  
DR InterPro; IPR001173; Glyco trans\_2.  
DR InterPro; IPR001951; Histone\_H4.  
DR Pfam; PF03142; Chitin\_synth\_2; 1.  
DR Pfam; PF00173; Cyt-b5; 1.  
DR PROSITE; PS00047; HISTONE\_H4; UNKNOWN 1.  
SQ SEQUENCE 1230 AA; 137504 MW; A8299B458361D870 CRC64;

Query Match 87.8%; Score 36; DB 2; Length 1230;  
Best Local Similarity 87.5%; Pred. No. 1.9e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QQYSSSPT 8  
|||  
Db 37 QQYSSPPT 44

RESULT 2  
Q6P0C7 PRELIMINARY; PRT; 230 AA.  
AC Q6P0C7;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein scamp5.  
GN Name=scamp5;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;

```
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marsden K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaeetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fayon J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Kryzyski M.I., Skalska U., Smalius D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC065670; AAH65670.1; -.
DR ZFIN; ZDB-GENE-030131-3188; scamp5.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR007273; SCAMP.
DR Pfam; PF04144; SCAMP; 1.
KW Hypothetical protein.
SQ SEQUENCE 230 AA; 25568 MW; C8080603A099B786 CRC64;

Query Match 85.4%; Score 35; DB 2; Length 230;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOYSSSPT 8
Db 216 QOYSAAPT 223

RESULT 3
Q8BJ42 PRELIMINARY; PRT; 1059 AA.
AC Q8BJ42;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Mus musculus adult male olfactory brain cDNA, RIKEN full-length
DE enriched library, clone:6430596N04 product:DISKS LARGE-ASSOCIATED
DE PROTEIN 2 (DAP-2) (SAP90/PSD-95-ASSOCIATED PROTEIN 2) (SAPAP2) (PSD-
DE 95/SAP90 BINDING PROTEIN 2) homolog.
GN Name=Dlgap2;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
```

```
RL Nature 409:685-690(2001).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RA The FANTOM Consortium;
RA "Analysis of the mouse transcriptome based on functional annotation of
RA 60,770 full-length cDNAs.";
RT Nature 420:563-573(2002).
RL [4]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayateu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imorani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai C., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AK032564; BAC27927.1; -.
DR MGD; MGI:2443181; Dlgap2.
DR GO; GO:0007267; P:cell-cell signaling; IEA.
DR InterPro; IPR005026; GKAP.
DR Pfam; PF03359; GKAP; 1.
SQ SEQUENCE 1059 AA; 119043 MW; 5DA511A902533F1D CRC64;

Query Match 82.9%; Score 34; DB 2; Length 1059;
Best Local Similarity 87.5%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QOYSSSPT 8
Db 58 QOYWSWPT 65

RESULT 4
DLF2_RAT STANDARD; PRT; 1089 AA.
ID DLF2_RAT
AC P97837;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Disks large-associated protein 2 (DAP-2) (SAP90/PSD-95-associated
DE protein 2) (SAPAP2) (PSD-95/SAP90 binding protein 2) (Fragment).
GN Name=Dlgap2; Synonyms=Dap2;
```

OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=97277335; PubMed=9115257; DOI=10.1074/jbc.272.18.11943;  
 RA Takeuchi M., Hata Y., Hirao K., Toyoda A., Irie M., Takai Y.;  
 RT "SAPAPs. A family of PSD-95/SAP90-associated proteins localized at  
 post-synaptic density.";  
 RL J. Biol. Chem. 272:11943-11951(1997).  
 CC -!- FUNCTION: May play a role in the molecular organization of  
 synapses and neuronal cell signaling. Could be an adapter protein  
 linking ion channel to the subsynaptic cytoskeleton. May induce  
 enrichment of PSD-95/SAP90 at the plasma membrane.  
 CC -!- SUBUNIT: Interacts with DLG1 and DLG4/PSD-95.  
 CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).  
 CC -!- TISSUE SPECIFICITY: Expressed in brain.  
 CC -!- SIMILARITY: Belongs to the SAPAP family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U67138; AAB48588.1; ALT\_INIT.  
 DR IntAct; P97837; -.  
 DR RGD; 620224; Dlgap2.  
 DR GO; GO:0005883; C:neurofilament; NAS.  
 DR GO; GO:0005815; F:protein binding; NAS.  
 DR GO; GO:0007270; P:nervous system synaptic transmission; NAS.  
 DR InterPro; IPR005026; GKAP.  
 DR Pfam; PF033359; GKAP; 1.  
 KW Membrane.  
 FT NON TER.  
 FT NON TER.  
 SQ SEQUENCE 1089 AA; 122233 MW; FEP8F8D2EB38B8A CRC64;  
 Query Match 82.9%; Score 34; DB 1; Length 1089;  
 Best Local Similarity 87.5%; Pred. No. 4.4e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QYSSSPT 8  
 DB 88 QYSSWSPT 95  
 RESULT 5  
 O41162  
 ID O41162 PRELIMINARY; PRT; 87 AA.  
 AC O41162;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE A680R protein.  
 GN Name=a680R;  
 OS Paramesitium bursaria chlorocella virus 1 (PBCV-1).  
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.  
 OX NCBI\_TaxID=10506;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98022962; PubMed=9356347; DOI=10.1006/viro.1997.8805;  
 RA Li Y., Lu Z., Sun L., Ropp S., Kutish G.F., Rock D.L., Van Etten J.L.;  
 RT "Analysis of 74 kb of DNA located at the right end of the 330-kb  
 chlorocella virus PBCV-1 genome.";  
 RL Virology 237:360-377(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20013326; PubMed=10544099; DOI=10.1006/viro.1999.9972;  
 RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,

RA Lisec A.D., Nickerson K.W., Van Etten J.L.;  
 RT "Chlorocella virus PBCV-1 encodes a functional homospermidine  
 synthase.";  
 RL Virology 263:254-262(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20478054; PubMed=11021991; DOI=10.1006/viro.2000.0500;  
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;  
 RT "Characterization of a beta-1,3-glucanase encoded by chlorocella virus  
 PBCV-1.";  
 RL Virology 276:27-36(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Gurnon J.R.;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U42580; AAC97053.1; -.  
 DR PIR; T18182; T18182.  
 SQ SEQUENCE 87 AA; 10337 MW; B6F28E923B153329 CRC64;  
 Query Match 80.5%; Score 33; DB 2; Length 87;  
 Best Local Similarity 75.0%; Pred. No. 42;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QYSSSPT 8  
 DB 77 QYSSKPT 84  
 RESULT 6  
 KV3D\_HUMAN  
 ID KV3D\_HUMAN STANDARD; PRT; 109 AA.  
 AC P01622;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Ig kappa chain V-III region Ti.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=72188439; PubMed=5027703;  
 RA Suter L., Barnikol H.U., Watanabe S., Hilschmann N.;  
 RT "Rule of antibody structure. The primary structure of a monoclonal  
 immunoglobulin L-chain of kappa-type, subgroup 3 (Bence-Jones protein  
 Ti). IV. The complete amino acid sequence and its significance for the  
 mechanism of antibody production.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 353:189-208(1972).  
 CC -!- MISCELLANEOUS: The C region of this chain has the INV (3) marker.  
 CC -!- MISCELLANEOUS: This is a Bence-Jones protein.  
 DR PIR; A01895; K3HUTI.  
 DR HSPP; P01625; ILVE.  
 DR GO; GO:0005576; C:extracellular; NAS.  
 DR GO; GO:0003823; F:antigen binding; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; ig; 1.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 DR Bence-Jones protein; Direct protein sequencing;  
 KW Immunoglobulin V region.  
 FT DISULFID 23 89 By similarity.  
 FT NON TER 109 109  
 SQ SEQUENCE 109 AA; 11788 MW; 8C35058CDD7749BC CRC64;  
 Query Match 80.5%; Score 33; DB 1; Length 109;  
 Best Local Similarity 75.0%; Pred. No. 54;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QYSSSPT 8  
 DB 77 QYSSKPT 84

```

Db          90 QOYGSPTS 97

RESULT 7
ID Q62954          PRELIMINARY;      PRT;    116 AA.
AC Q62954;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE EGF receptor (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=97184212; PubMed=9030624;
RA Carroll S.L., Miller M.L., Frohnert P.W., Kim S.S., Corbett J.A.;
RT "Expression of neurogins and their putative receptors, ErbB2 and
RT ErbB3, is induced during Wallerian degeneration.";
RL J. Neurosci. 17:1642-1659 (1997).
DR EMBL; U52529; AAC53049.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_TER          1
FT NON_TER          116
SQ SEQUENCE 116 AA; 12678 MW; 90A0194257C5F046 CRC64;

Query Match      80.5%; Score 33; DB 2; Length 116;
Best Local Similarity 75.0%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYSSSPT 8
|:|||||
Db 36 QRYSSDPT 43

RESULT 8
KV3K_HUMAN
ID KV3K_HUMAN      STANDARD;      PRT;    128 AA.
AC P06311;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1998 (Rel. 38, Last annotation update)
DE Ig kappa chain V-III region IARC/BL41 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86041852; PubMed=2997711;
RA Klobeck H.G., Meindl A., Combratio G., Solomon A., Zachau H.G.;
RT "Human immunoglobulin kappa light chain genes of subgroups II and
RT III.";
RL Nucleic Acids Res. 13:6499-6513 (1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z00021; CAA77316.1; -.
DR PIR; A01899; K3HU41.
DR HSSP; P01625; 1E9Q.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.

```

---

```

DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL          1
FT CHAIN           21
FT CHAIN           128
FT DOMAIN           21
FT DOMAIN           43
FT DOMAIN           44
FT DOMAIN           54
FT DOMAIN           55
FT DOMAIN           69
FT DOMAIN           70
FT DOMAIN           76
FT DOMAIN           77
FT DOMAIN           108
FT DOMAIN           109
FT DOMAIN           117
FT DOMAIN           118
FT DISULFID         43
FT NON_TER          128
FT NON_TER          128
SQ SEQUENCE 128 AA; 14070 MW; CC89570FE3B9012 CRC64;

Query Match      80.5%; Score 33; DB 1; Length 128;
Best Local Similarity 85.7%; Pred. No. 65;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOYSSSPT 7
|:|||||
Db 109 QOYSTSP 115

RESULT 9
Q811C3
ID Q811C3          PRELIMINARY;      PRT;    131 AA.
AC Q811C3;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Immunoglobulin gamma-3 kappa chain precursor (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MRL/MPJ-lpr/lpr; TISSUE=Spleen;
RX MEDLINE=93156722; PubMed=8429833; DOI=10.1016/0161-5890(93)90089-T;
RA Takahashi S., Itoh J., Nose M., Ono M., Yamamoto T., Kyogoku M.;
RT "Cloning and cDNA sequence analysis of nephritogenic monoclonal
RT antibodies derived from an MRL/lpr lupus mouse.";
RL Mol. Immunol. 30:177-182 (1993).
DR EMBL; D14629; BAA03482.1; -.
DR HSSP; P01679; 2FEJ.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
KW Signal.
FT SIGNAL          1
FT CHAIN           23
FT CHAIN           >131
FT NON_TER          131
FT NON_TER          131
SQ SEQUENCE 131 AA; 14083 MW; 5E83656954666E9E CRC64;

Query Match      80.5%; Score 33; DB 2; Length 131;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QOYSSSPT 8
|:|||||
Db 112 QOYDSSPS 119

RESULT 10
Q6NGH3
ID Q6NGH3          PRELIMINARY;      PRT;    249 AA.
AC Q6NGH3;

```



DT	05-JUL-2004	(TREMBlrel. 27, Created)	CC	(class III) peroxidase subfamily.
DT	05-JUL-2004	(TREMBlrel. 27, Last sequence update)	DR	EMBL; U12314; AAA20472.1; -.
DT	05-JUL-2004	(TREMBlrel. 27, Last annotation update)	DR	HSSP; P22195; 1SCH.
DE	Ribonuclease III	(EC 3.1.26.3)	DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
GN	Name=rnc; Ordered accession=DIP1544;		DR	GO; GO:0004601; F:peroxidase activity; IEA.
OS	Corynebacterium diphtheriae.		DR	GO; GO:0006979; P:response to oxidative stress; IEA.
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		DR	InterPro; IPR002016; Peroxidase_super.
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.		DR	InterPro; IPR010255; Peroxidase_super.
OX	NCBI_TaxID=1717;		DR	InterPro; IPR000823; Plant_peroxidase.
RN	[1]		DR	Pfam; PF00141; peroxidase; 1.
RP	SEQUENCE FROM N.A.		DR	PRINTS; PR00458; PEROXIDASE.
RC	STRAIN=Biotype Gravis / NCTC 13129;		DR	PRINTS; PR00461; PLPEROXIDASE.
RX	MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;		DR	PROSITE; PS00436; PEROXIDASE_2; 1.
RA	Cerdeno-Tarraga A.-M., Bestratiou A., Dover L.G., Holden M.T.G.,		DR	PROSITE; PS00873; PEROXIDASE_4; 1.
RA	Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,		KW	Heme; Iron; Oxidoreductase; Peroxidase; Signal.
RA	De Zoyea A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,		FT	SIGNAL 1 23 Potential.
RA	Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,		FT	CHAIN 24 307 peroxidase.
RA	Rabinowitch E., Rutherford K.M., Thomson N.R., Unwin L.,		SQ	SEQUENCE 307 AA; 32508 MW; 526B44B3A7C75FA7 CRC64;
RT	Whitehead S., Barrell B.G., Parkhill J.;			
RT	"The complete genome sequence and analysis of Corynebacterium			
RT	diphtheriae NCTC13129."			
RL	Nucleic Acids Res. 31:6516-6523(2003).			
DR	EMBL; EX248358; CAE50069.1; -.			
DR	GO; GO:0003622; C:intracellular; IEA.			
DR	GO; GO:0003725; F:double-stranded RNA binding; IEA.			
DR	GO; GO:0016787; F:hydrolase activity; IEA.			
DR	GO; GO:0004525; F:ribonuclease III activity; IEA.			
DR	GO; GO:0006396; P:RNA processing; IEA.			
DR	InterPro; IPR001159; DS_RBD.			
DR	InterPro; IPR000999; RNase_III.			
DR	Pfam; PF00035; dsrm; 1.			
DR	SMART; SM00358; Ribonuclease_3; 1.			
DR	SMART; SM00535; RIBOC; 1.			
DR	PROSITE; PS00137; DS_RBD; 1.			
DR	PROSITE; PS00517; RNase_3_1; 1.			
DR	PROSITE; PS0142; RNase_3_2; 1.			
KW	Complete proteome; Hydrolase.			
SQ	SEQUENCE 249 AA; 27353 MW; 825E51EDF5F2279 CRC64;			
	Query Match 80.5%; Score 33; DB 2; Length 249;			
	Best Local Similarity 85.7%; Pred. No. 1.4e+02;			
	Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
QY	1 QOYSSSP 7			
Db	79 QOYTSSP 85			
	RESULT 11			
ID	Q43416	PRELIMINARY; PRT; 307 AA.		
AC	Q43416			
DT	01-NOV-1996	(TREMBlrel. 01, Created)		
DT	01-NOV-1996	(TREMBlrel. 01, Last sequence update)		
DT	01-MAR-2004	(TREMBlrel. 26, Last annotation update)		
DE	Peroxidase precursor (EC 1.11.1.7).			
OS	Cenchrus ciliaris (Buffelgrass) (Pennisetum ciliare).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	PACCAD clade; Panicoideae; Paniceae; Cenchrus.			
OX	NCBI_TaxID=35872;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Wounded and unwounded leaves and stems;			
RA	Ross A.H.;			
RA	"Investigation of peroxidase genes and genetic transformation in			
RT	buffel grass."			
RL	Thesis (1994), Botany, University of Queensland.			
CC	-1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.			
CC	-1- COFACTOR: Binds 1 heme B (iron-protophyrin IX) group and 2			
CC	calcium ions per subunit (By similarity).			
CC	-1- SIMILARITY: Belongs to the peroxidase family. Classical plant			

Query Match	80.5%;	Score 33;	DB 2;	Length 327;
Best Local Similarity	85.7%;	Pred. No. 1.8e+02;		
Matches	6;	Conservative	1;	Mismatches 0; Indels 0; Gaps 0;
QY	1 QOYSSSP 7			
Db	284 QOYSSNP 290			
	RESULT 12			
Q6ER51		PRELIMINARY; PRT; 327 AA.		
ID	Q6ER51			
AC	Q6ER51			
DT	25-OCT-2004	(TREMBlrel. 28, Created)		
DT	25-OCT-2004	(TREMBlrel. 28, Last sequence update)		
DT	25-OCT-2004	(TREMBlrel. 28, Last annotation update)		
DE	Putative bacterial-induced peroxidase.			
GN	Names=OSUNBa0082C09.13;			
OS	Oryza sativa (japonica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=39947;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Sasaki T., Matsumoto T., Katayose Y.;			
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.			
CC	-1- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.			
CC	-1- COFACTOR: Binds 1 heme B (iron-protophyrin IX) group and 2			
CC	calcium ions per subunit (By similarity).			
CC	-1- SIMILARITY: Belongs to the peroxidase family. Classical plant			
CC	(class III) peroxidase subfamily.			
DR	EMBL; AP005613; BAD28869.1; -.			
DR	GO; GO:0004601; F:peroxidase activity; IEA.			
DR	GO; GO:0006979; P:response to oxidative stress; IEA.			
DR	InterPro; IPR002016; Peroxidase.			
DR	InterPro; IPR010255; Peroxidase_super.			
DR	InterPro; IPR000823; Plant_peroxidase.			
DR	Pfam; PF00141; peroxidase; 1.			
DR	PRINTS; PR00458; PEROXIDASE.			
DR	PRINTS; PR00461; PLPEROXIDASE.			
DR	PROSITE; PS00436; PEROXIDASE_1; UNKNOWN_1.			
DR	PROSITE; PS00436; PEROXIDASE_2; 1.			
DR	PROSITE; PS00873; PEROXIDASE_4; 1.			
KW	Heme; Iron; Oxidoreductase; Peroxidase.			
SQ	SEQUENCE 327 AA; 34427 MW; B26DF62A9B39CE94 CRC64;			

```

RESULT 13
Q8MV41 PRELIMINARY; PRT; 339 AA.
ID Q8MV41
AC Q8MV41
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Relish R6 isoform (Fragment).
GN Name=relish;
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
OX NCBI_TaxID=7159;
RN [1]_TaxID=7159;
RN SEQUENCE FROM N.A.
RP MEDLINE=22133804; PubMed=12119421; DOI=10.1073/pnas.162345999;
RX Shin S.W., Kokoza V., Ahmed A., Raikhel A.S.;
RT "Characterization of three alternatively spliced isoforms of the
RT Rel/NF-kappa B transcription factor Relish from the mosquito Aedes
RT aegypti.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9978-9983(2002).
[2]
RN SEQUENCE FROM N.A.
RP Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RL Shin S.W., Raikhel A.;
DR EMBL; AF499466; AAM97897.1; -.
DR HSSP; Q04207; IMY7.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR011539; RHD.
DR Pfam; PF00554; RHD; 1.
DR PROSITE; PS50254; REL_2; 1.
FT NON_TER 1
FT NON_TER 339
SQ SEQUENCE 339 AA; 38019 MW; 38F8BC07387D8F2C CRC64;

Query Match 80.5%; Score 33; DB 2; Length 339;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QOYSSSPT 8
Db 12 QOLSSSPT 19

RESULT 14
Q75MF2 PRELIMINARY; PRT; 387 AA.
ID Q75MF2
AC Q75MF2;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein EGFR (Fragment).
GN Name=EGFR;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RN SEQUENCE FROM N.A.
RP MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
RX Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
RA Wyllie K., Sekhon N., Becker M.C., O'Laughlin M.D., Schaller M.E.,
RA Fewell G.A., Delhaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
RA Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
RA Ozersky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
RA Kozlowicz-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
RA Tin-Wollam A.M., Abbott A., Minx P., Maupin R., Strommatt C.,
RA Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,

```

```

RA Wendl M.C., Yang S.P., Schultz B.R., Wallis J.W., Spieth J.,
RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
RA Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
RA Clifton S.W., Chissee S.L., Marra M.A., Raymond C., Haugen E.,
RA Gillett W., Zhou Y., James R., Phelps K., Iadanoto S., Bubbs K.,
RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
RA Baertsch R.A., Portnoy M.E., Keibler E., Flicek P., Bork P., Suyama M.,
RA Bailey J.A., Torrents D., Chinwalla A.T., Gish W.R.,
RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
RA Waterston R.H., Wilson R.K.;
RT "The DNA sequence of human chromosome 7.";
RL Nature 424:157-164(2003).
[2]
RN SEQUENCE FROM N.A.
RP Waterston R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]_TaxID=7159;
RN SEQUENCE FROM N.A.
RP Wilson R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC073324; AAS07524.1; -.
DR HSSP; P11362; IAGW.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 387
SQ SEQUENCE 387 AA; 43532 MW; 196782B43B417155 CRC64;

Query Match 80.5%; Score 33; DB 2; Length 387;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QOYSSSPT 8
Db 244 QRYSSDPT 251

RESULT 15
Q6LHL4 PRELIMINARY; PRT; 509 AA.
ID Q6LHL4
AC Q6LHL4;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein VVA1385.
GN Name=VVA1385; OrderedLocustNames=BBPRB1345;
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]_TaxID=74109;
RN SEQUENCE FROM N.A.
RP Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
RA Cestaro A., Malacrida G., Simonati B., Cannata N., Bartlett D.,
RA Valle G.;
RT "Genome analysis of Photobacterium profundum reveals the complexity of
RT high pressure adaptations.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR378679; CAG23216.1; -.
KW Complete proteome; Hypothetical protein.

```

SQ SEQUENCE 509 AA; 57807 MW; EA965D94F13CE979 CRC64;  
Query Match 80.5%; Score 33; DB 2; Length 509;  
Best Local Similarity 75.0%; Pred. No. 3e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QQYSSSPT 8  
Db 44 QQFSSSPS 51

Search completed: August 9, 2005, 14:19:38  
Job time : 87.8462 secs

This Page Blank (nsptot)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 03:48:39 ; Search time 3308.69 Seconds  
(without alignments)  
5184.273 Million cell updates/sec

Title: US-10-089-452-1  
Perfect score: 354  
Sequence: 1 gaggtgcagctgctcgagca.....ccactctcacagttctctca 354

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sv.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	354	100.0	354	6	AX113426 Sequence
2	354	100.0	354	6	AX113561 Sequence
3	354	100.0	354	6	AX816228 Sequence
4	272.2	76.9	747	12	AY270182 Synthetic
5	271.4	76.7	348	10	AF253061 Mus muscu
6	268.2	75.8	406	10	M28251 Mouse Ig re
7	268.2	75.8	443	6	I09505 Sequence 3
8	268	75.7	333	10	AF218695 Mus muscu
9	268	75.7	411	10	AF045488 Mus muscu
10	267.2	75.5	719	6	AR100151 Sequence
11	267.2	75.5	720	6	AR117135 Sequence
12	267.2	75.5	1797	6	I84705 Sequence 2
13	266.6	75.3	470	10	M19898 Mouse Ig re
14	266.6	75.3	478	10	MUSIGHAAN
15	266.6	75.3	783	10	AB007986
16	264.8	74.8	330	10	AF218693 Mus muscu
17	264.2	74.6	351	6	E13088 cDNA encodi
18	264.2	74.6	354	6	AR026097 Sequence
19	263.2	74.4	333	10	AF218648 Mus muscu

20	261.6	73.9	330	10	AF218674
21	261.2	73.8	654	10	U00940 Mus musculu
22	261.2	73.8	732	12	XX044796
23	261.2	73.8	898	6	BD096937
24	261.2	73.8	1794	6	BD096938 Mutated O
25	261.2	73.8	1817	6	AX739887 Sequence
26	261.2	73.8	6091	6	C0789322 Sequence
27	261.2	73.8	6844	6	AX739888 Sequence
28	260.2	73.5	351	6	E04482 DNA encodin
29	260.2	73.5	418	10	D82081 Mus musculu
30	260.2	73.5	1568	6	E33134 Humanized a
31	260.2	73.5	1570	6	A22261 M.musculus
32	260.2	73.5	1570	6	A77138 Sequence 6
33	260.2	73.5	1570	6	AR559698 Sequence
34	260.2	73.5	10511	6	AX666316 Sequence
35	259.6	73.3	1653	6	BD206135 Polyvalen
36	259.6	73.3	1653	6	AX011208 Sequence
37	259.6	73.3	1698	6	BD206134 Polyvalen
38	259.6	73.3	1698	6	AX011206 Sequence
39	259.6	73.3	1817	6	AX739885 Sequence
40	258.8	73.1	318	10	AY171930 Mus muscu
41	258.6	73.1	1570	6	AR029102 Sequence
42	258.6	73.1	1570	6	AR409372 Sequence
43	258.4	73.0	333	10	AF218700 Mus muscu
44	258.4	73.0	354	10	MUSIGHAEF
45	258	72.9	642	6	AX014731 Sequence

ALIGNMENTS

RESULT 1  
AX113426  
LOCUS AX113426 354 bp DNA linear PAT 01-MAY-2001  
DEFINITION Sequence 1 from Patent WO0127612.  
ACCESSION AX113426  
VERSION AX113426.1 GI:13939694  
KEYWORDS Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Reiter,C., Cullmann,G., Lakner,M., Truee,A., Dehnert,S. and Schwartz,G.  
TITLE Immuno-chromatographic rapid assay in order to detect acid-resistant microorganisms in the stool  
JOURNAL Patent: WO 0127612-A 1 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung  
mbH (DE) Location/Qualifiers  
source 1..354  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"

ORIGIN  
Query Match 100.0%; Score 354; DB 6; Length 354;  
Best Local Similarity 100.0%; Pred. No. 7e-105;  
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAGGTGCAGCTGCTCGAGCAGCCTGGGGCTGAACTGGGCAAAACCTGGGGCCCTCAGTGAAG 60  
DB 1 GAGGTGCAGCTGCTCGAGCAGCCTGGGGCTGAACTGGGCAAAACCTGGGGCCCTCAGTGAAG 60  
QY 61 ATGCTCTCAAGGCTTCTGGCTACACCTTTACTACTACTGATTCTCCTGGGTGAAACAG 120  
DB 61 ATGCTCTCAAGGCTTCTGGCTACACCTTTACTACTACTGATTCTCCTGGGTGAAACAG 120  
QY 121 AGGCTGCACAGGGTCTGAAATGGATTGGATACATTAATCTGCCACCTGGTTCACCTTCT 180  
DB 121 AGGCTGCACAGGGTCTGAAATGGATTGGATACATTAATCTGCCACCTGGTTCACCTTCT 180

```
Qy 181 TACAATCAGGACTTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCTCCACACAGCC 240
Db 181 TACAATCAGGACTTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCTCCACACAGCC 240

Qy 241 TACATGCAGCTGACCGAGCTGACATCTCAGAGACTCTTCAGTCTATTACTGTGCAAGAGAG 300
Db 241 TACATGCAGCTGACCGAGCTGACATCTCAGAGACTCTTCAGTCTATTACTGTGCAAGAGAG 300

Qy 301 GGGTACGACGGGTTTGACTCTCTGGGGCCAAAGCAGCACTCTCACAGTCTCCTCA 354
Db 301 GGGTACGACGGGTTTGACTCTCTGGGGCCAAAGCAGCACTCTCACAGTCTCCTCA 354

RESULT 2
LOCUS AX113561 354 bp DNA linear PAT 01-MAY-2001
DEFINITION Sequence 1 from Patent WO0127613.
ACCESSION AX113561
VERSION AX113561.1 GI:13939765
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Reiter,C., Cullmann,G., Heppner,P., Ringeis,A., Mueller,H. and
Haindl,E.
TITLE Improved method for the detection of acid resistant microorganisms
in a stool
JOURNAL Patent: WO 0127613-A 1 19-APR-2001;
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung
(DE)
FEATURES
source Location/Qualifiers
1. 354
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 100.0%; Score 354; DB 6; Length 354;
Best Local Similarity 100.0%; Pred. No. 7e-105;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGTGCAGCTGCTCGAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCCTCAGTGAAG 60
Db 1 GAGGTGCAGCTGCTCGAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCCTCAGTGAAG 60

Qy 61 ATGTCCTGCAAGGCTTCTGGCTACACCTTTACTAATCTACTGGATTCTACTGGGTGAAACAG 120
Db 61 ATGTCCTGCAAGGCTTCTGGCTACACCTTTACTAATCTACTGGATTCTACTGGGTGAAACAG 120

Qy 121 AGGCTCGACAGGCTCTGAAATGGATTGGATACATTAATCTGCGACTGGTTCCACTTCT 180
Db 121 AGGCTCGACAGGCTCTGAAATGGATTGGATACATTAATCTGCGACTGGTTCCACTTCT 180

Qy 181 TACAATCAGGACTTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCTCCACACAGCC 240
Db 181 TACAATCAGGACTTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCTCCACACAGCC 240

Qy 241 TACATGCAGCTGACCGAGCTGACATCTGAGAGCTCTTCAGTCTATTACTGTGCAAGAGAG 300
Db 241 TACATGCAGCTGACCGAGCTGACATCTGAGAGCTCTTCAGTCTATTACTGTGCAAGAGAG 300

Qy 301 GGGTACGACGGGTTTGACTCTCTGGGGCCAAAGCAGCACTCTCACAGTCTCCTCA 354
Db 301 GGGTACGACGGGTTTGACTCTCTGGGGCCAAAGCAGCACTCTCACAGTCTCCTCA 354

RESULT 3
LOCUS AX816228 354 bp DNA linear PAT 09-DEC-2003
DEFINITION Sequence 1 from Patent EP1336850.
ACCESSION AX816228
```

```
VERSION AX816228.1 GI:39646759
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Reiter,C., Cullmann,G., Mueller,H., Heppner,P., Haindl,E. and
Ringeis,A.
TITLE Improved method for the detection of acid resistant microorganisms
in a stool
JOURNAL Patent: BP 1336850-A 1 20-AUG-2003;
Connex Gesellschaft zur Optimierung von Forschung und Ent wicklung
(DE)
FEATURES
source Location/Qualifiers
1. 354
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 100.0%; Score 354; DB 6; Length 354;
Best Local Similarity 100.0%; Pred. NO. 7e-105;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGTGCAGCTGCTCGAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCCTCAGTGAAG 60
Db 1 GAGGTGCAGCTGCTCGAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCCTCAGTGAAG 60

Qy 61 ATGTCCTGCAAGGCTTCTGGCTACACCTTTACTAATCTACTGGATTCTACTGGGTGAAACAG 120
Db 61 ATGTCCTGCAAGGCTTCTGGCTACACCTTTACTAATCTACTGGATTCTACTGGGTGAAACAG 120

Qy 121 AGGCTCGACAGGCTCTGAAATGGATTGGATACATTAATCTGCGACTGGTTCCACTTCT 180
Db 121 AGGCTCGACAGGCTCTGAAATGGATTGGATACATTAATCTGCGACTGGTTCCACTTCT 180

Qy 181 TACAATCAGGACTTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCTCCACACAGCC 240
Db 181 TACAATCAGGACTTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCTCCACACAGCC 240

Qy 241 TACATGCAGCTGACCGAGCTGACATCTGAGAGCTCTTCAGTCTATTACTGTGCAAGAGAG 300
Db 241 TACATGCAGCTGACCGAGCTGACATCTGAGAGCTCTTCAGTCTATTACTGTGCAAGAGAG 300

Qy 301 GGGTACGACGGGTTTGACTCTCTGGGGCCAAAGCAGCACTCTCACAGTCTCCTCA 354
Db 301 GGGTACGACGGGTTTGACTCTCTGGGGCCAAAGCAGCACTCTCACAGTCTCCTCA 354

RESULT 4
LOCUS AY270182 747 bp DNA linear SYN 05-MAY-2003
DEFINITION Synthetic construct anti-MRPI single-chain antibody fragment A5
gene, partial cds.
ACCESSION AY270182
VERSION AY270182.1 GI:30385687
KEYWORDS
SOURCE Synthetic construct
ORGANISM synthetic construct
synthetic sequences; artificial sequences.
REFERENCE 1 (bases 1 to 747)
AUTHORS Binyamin,L., Asgaraf,Y.G. and Reiter,Y.
TITLE Isolation and characterization of a small recombinant single-chain
Fv antibody fragment targeted to an extracellular epitope of
multi-drug resistant protein MRPI
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 747)
AUTHORS Binyamin,L. and Reiter,Y.
TITLE Direct Submission
JOURNAL Submitted (07-APR-2003) Biology, Technion, Haifa 32000, Israel
FEATURES
source Location/Qualifiers
1. 747
```

```
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
1..>747
/feature="targeted to an extracellular domain"
/transl_table=1
/codon_start=1
/product="anti-MRPI single-chain antibody fragment A5"
/protein_id="AAP23214.1"
/db_xref="GI:30385688"
/translation="MAQVLOQSGAELARPGASVMSCKASGYFTFTSYTMHWKORPG
QGLEWIGYINPSSGYTNQKFKDKATLADKSSATYMLSSLTSDSAVYCARGN
YRFDYWGQGTTLTVSSGGSGGGGGSDVMTQPLSLPVLGDQASISCRSS
QSLVHSNGNTVLHWYLOKPGOSPKLLIVKSNRPSGVDPDRFSGSGSGTFTLKISRVE
AEDLGVYVCFQGLHVPFTFGSGTKLEIKRAA"

CDS
FEATURES
source
Location/Qualifiers
1..348
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Biozzi"
/db_xref="taxon:10090"
/cell_line="9A8"
/cell_type="hybridoma"
<1..>348
/feature="anti-idiotypic catalytic antibody"
/codon_start=1
/product="IgG heavy chain variable region"
/protein_id="AAF70264.1"
/db_xref="GI:7839401"
/translation="OVLOQSGAELAKPGASVKLSCKASGYFTFTSYTMHWKORPGQ
LEWIGYINPSSGYTNQKFKDKATLADKSSATYMLSSLTSDSAVYCARSYYG
PDYWGQGTTLTVSS"

ORIGIN
Query Match 76.9%; Score 272.2; DB 12; Length 747;
Best Local Similarity 87.4%; Pred. No. 6.2e-78;
Matches 298; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 13 CTCGAGCAGCTGGGGCTGAACCTGCAAAACCTGGGGCTCAGTGAAGATGTCCTGCAAG 72
Db 16 CTCGAGCAGTCTGGGGCTGAACCTGCAAGAGCCTGGGGCTCAGTGAAGATGTCCTGCAAG 75
QY 73 GCTTCTGGCTACACCTTTTACTAACTTACTGGATTCACTGGGTGAACACAGAGCGCTGACAG 132
Db 76 GCTTCTGGCTACACCTTTTACTAGCTACAGATGCACTGGGTGAACACAGAGCGCTGCAAG 135
QY 133 GCTTGAATGATTTGGATACATTAACTCTGCACTGGTTCACCTTCTTCAATCAGGAC 192
Db 136 GGTCTGGAATGGATTGGATACATTAACTCTAGCAGTGGTTATTAATCAATCAGAAG 195
QY 193 TTTCAAGCAGGCGCATTGACCGCAGCAAGTCTCTCAGCAGCAGCTCATGCAAGCTG 252
Db 196 TTTCAAGCAGGCGCATTGACCTGAGCAAAATCTCAGCAGCAGCTCATGCAAGCTG 255
QY 253 ACCAGCTGACATCTGAGGACTCTGAGTCTTACTGTTACTGTCGAAGAGGGGTACGACGGG 312
Db 256 AGCAGCTGACATCTGAGGACTCTGAGTCTTACTGTCGAAGAGGGAATATAGGTAC 315
QY 313 TTTGACTCTCTGGGGCAAGGACCACTCTCAGAGTCTCTC 353
Db 316 TTTGACTCTCTGGGGCAAGGACCACTCTCAGGCTCTCTC 356

RESULT 5
AF253061 348 bp mRNA linear ROD 06-DEC-2000
LOCUS
DEFINITION
Mus musculus IgG heavy chain variable region mRNA, partial cds.
ACCESSION
AF253061.1 GI:7839400
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 348)
Demin.A.V., Titov.M.I., Avalle.B., Tramontano.A., Paul.S.,
Kolesnikov.A.V., Kozlov.A.V., Alexandrova.E.S., Koralewski.F.,
Thomas.D., Gabibov.A.G. and Friboulet.A.
Enzyme mimicry by the anti-idiotypic antibody approach
Proc. Natl. Acad. Sci. U.S.A. 97 (25), 13526-13531 (2000)
20558540
11095704
REFERENCE
2 (bases 1 to 348)
Kolesnikov.A.V., Kozlov.A.V., Alexandrova.E.S., Koralewski.F.,
Demin.A.V., Titov.M.I., Avalle.B., Tramontano.A., Paul.S.,
Thomas.D., Gabibov.A.G. and Friboulet.A.
Direct Submission
Submitted (06-APR-2000) Laboratory of Biocatalysis, Shemyakin &
Ovchinnikov Institute of Bioorganic Chemistry, Mikluho-Maklaya
16/10, Moscow 117871, Russia

RESULT 6
MUSIGHMX 406 bp mRNA linear ROD 12-FEB-2001
LOCUS
DEFINITION
Mouse Ig rearranged gamma-chain (G-2a) mRNA V-region (V-J2),
partial cds.
ACCESSION
M28251.1 GI:195663
VERSION
KEYWORDS
constant region; immunoglobulin gamma-chain; immunoglobulin heavy
chain; joining region; processed gene; variable region.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 406)
Queen.C.
Unpublished
2 (sites)
REFERENCE
AUTHORS
Queen.C., Schneider.W.P., Selick.H.E., Payne.P.W., Landolfi.N.F.,
Duncan.J.F., Avdalovic.N.M., Levitt.M., Jungmans.R.P. and
Waldmann.T.A.
A humanized antibody that binds to the interleukin 2 receptor
Proc. Natl. Acad. Sci. U.S.A. 86 (24), 10029-10033 (1989)
90099290
2513570
PUBMED
COMMENT
Draft entry and computer-readable sequence kindly submitted by
```

```

FEATURES             C.Queen, 22-SEP-1989.
    source            Location/Qualifiers
        1..406
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="BALB/c"
        /db_xref="taxon:10090"
        /cell_type="Anti-Tac hybridoma"
        1..>406
        /note="Ig gamma 2a-chain V-region; (V-J2) precursor"
        /codon_start=1
        /protein_id="AA03013.2"
        /db_xref="GI:12750776"
        /translation="MERHWIPLFLFSVTAGVHSQVQLQDSGAELAKPQASVYKMSCKAS
        GYTSIRMHVVKORPQGLEWIGYINPSTGYEYNQKDKATLTADKSSSTAYMQL
        SSILTFEDSAVYCYARGGVFDYWGQGTTLTVSS"
        1..60
        /note="Ig gamma-chain signal peptide"
        61..>406
        /product="Ig gamma chain"
        351..352
        /note="V-region end/N-region start"
        362..363
        /note="N-region end/J2-region start"

    CDS
        10 CTGCTCAGCAGCCTGGGGCTGAACTGGCAAAACCTGGGGCCTCAGTGAAGATGTCCTGC 69
        64 CAGCTTCAGCAGTCTGGGGCTGAACTGGCAAAACCTGGGGCCTCAGTGAAGATGTCCTGC 123
        70 AAGCTTCTGGCTACACCTTTACTAACTACTGGATTCACTGGGGAACACAGAGCCTGGA 129
        124 AAGCTTCTGGCTACACCTTTACTAGCTACAGGATGCATGGGTAACACAGAGCCTGGA 183
        130 CAGGCTCGAAATGGATTGGATACATTAATCTCCCACTGGTTCCTCTTACAATCAG 189
        184 CAGGCTCGGAATGGAATGGATATATTAACTCTAGCACTGGGTATCTGAATACATCAG 243
        190 GACTTTCAGCAGAGGCGCACTTTGACCGCAGACAAGTCTCTCCACACAGCCTCATGCGAG 249
        244 AAGTTCAAGCAACAGGCGCACTTACTGCGACACAAATCTCCAGCAGACGCTCATGCAAA 303
        250 CTGACCAAGCCTGACATCTGAGGACTCTTCAGTCTTATTACTGTGCAAGAGGGGTACGAC 309
        304 CTGAGCAGCCTGACATTTGAGGACTCTGCACTCTATTACTGTGCAAGAGGGGGG---GGG 360
        310 GGGTTTGACTCTGGGGCCCAAGGCACCACTCTCACAGTCTCCTCA 354
        361 GTCTTTGACTACTGGGGCCCAAGGAACCACTCTCACAGTCTCCTCA 405

RESULT 7
LOCUS             I09505
DEFINITION        Sequence 3 from Patent WO 8909622.
ACCESSION         I09505
VERSION           I09505.1 GI:587789
KEYWORDS          Unknown.
SOURCE            Unknown.
ORGANISM          Unclassified.
REFERENCE          1 (bases 1 to 443)
AUTHORS           Queen,C.L.
TITLE             IL-2 RECEPTOR-SPECIFIC CHIMERIC ANTIBODIES
JOURNAL           Patent: WO 8909622-A 3 19-OCT-1989;
FEATURES          Location/Qualifiers
    source        1..443
                  /organism="unknown"
                  /mol_type="unassigned DNA"

```

```

ORIGIN
    Query Match      75.8%; Score 268.2; DB 6; Length 443;
    Best Local Similarity 88.1%; Pred. No. 1.2e-76;
    Matches 304; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

Qy 10 CTGCTCAGCAGCCTGGGGCTGAACTGGCAAAACCTGGGGCCTCAGTGAAGATGTCCTGC 69
Db 101 CAGCTTCAGCAGTCTGGGGCTGAACTGGCAAAACCTGGGGCCTCAGTGAAGATGTCCTGC 160

Qy 70 AAGCTTCTGGCTACACCTTTACTAACTACTGGATTCACTGGGGAACACAGAGCCTGGA 129
Db 161 AAGCTTCTGGCTACACCTTTACTAGCTACAGGATGCATGGGTAACACAGAGCCTGGA 220

Qy 130 CAGGCTCGAAATGGATTGGATACATTAATCTCCCACTGGTTCCTCTTACAATCAG 189
Db 221 CAGGCTCGGAATGGAATGGATATATTAACTCTAGCACTGGGTATCTGAATACATCAG 280

Qy 190 GACTTTCAGCAGAGGCGCACTTTGACCGCAGACAAGTCTCTCCACACAGCCTCATGCGAG 249
Db 281 AAGTTCAAGCAACAGGCGCACTTACTGCGACACAAATCTCCAGCAGACGCTCATGCAAA 340

Qy 250 CTGACCAAGCCTGACATCTGAGGACTCTTCAGTCTTATTACTGTGCAAGAGGGGTACGAC 309
Db 341 CTGAGCAGCCTGACATTTGAGGACTCTGCACTCTATTACTGTGCAAGAGGGGGG---GGG 397

Qy 310 GGGTTTGACTCTGGGGCCCAAGGCACCACTCTCACAGTCTCCTCA 354
Db 398 GTCTTTGACTACTGGGGCCCAAGGAACCACTCTCACAGTCTCCTCA 442

RESULT 8
LOCUS             AF218695
DEFINITION        Mus musculus clone nMev171 immunoglobulin heavy chain variable
                  region mRNA, partial cds.
ACCESSION         AF218695
VERSION           AF218695.1 GI:7381591
KEYWORDS          Mus musculus (house mouse)
SOURCE            Mus musculus
ORGANISM          Mus musculus
REFERENCE          1 (bases 1 to 333)
AUTHORS           Cassidy,R.L. and Kaushik,A.
TITLE             Influence of increased negative selection on VDJ rearrangements in
                  B-cells arising early in ontogeny
JOURNAL           Unpublished
REFERENCE          2 (bases 1 to 333)
AUTHORS           Cassidy,R.L. and Kaushik,A.
TITLE             Direct Submission
JOURNAL           Submitted (22-DEC-1999) Pathobiology, University of Guelph, Guelph,
                  Ontario N1G 2W1, Canada
FEATURES          Location/Qualifiers
    source        1..333
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /db_xref="taxon:10090"
                  /clone="nMev171"
                  /tissue_type="spleen, liver"
                  /dev_stage="neonatal"
                  /note="viable motheaten"
                  <1..>333
                  /notes="J558 family VDJ rearrangement; aa 7-114, kabat
                  numbering"
                  /codon_start=1
                  /product="immunoglobulin heavy chain variable region"
                  /protein_id="AAF61556.1"
                  /db_xref="GI:7381592"
                  /translation="SGAELAKPRASVKLSCKASGYTFTSYMHWVKORPQGLEWIG
                  INPSSGYTKYNQKFKDKATLTADKSSSTAAMQLSSLTYEDSAVYCAIYYDYDFDYG
                  QGTTLTVSS"

    CDS

```

ORIGIN





<i>Qy</i>	190	GACTTTCAGGACAGGGCCACTTTCACCGCAGACAAGTCTCTCCACCACAGCCTACATGCAG	249
<i>Db</i>	193	AAGTTCAAGGACAGGCCACATTGACTGCAGACAATCTCTCAGCACAGCCTACATGCAA	252
<i>Qy</i>	250	CTGACACAGCCTGACATCTGAGGACTCTTCAGTCTTATTACTGTGCAAGAGAGGGGTACGCAC	309
<i>Db</i>	253	CTGAGCACAGCCTGACATTTGAGGACTCTGCAGTCTTATTACTGTGCAAGAGAGGGGG--GGG	309
<i>Qy</i>	310	GGGTTTGACTCTCGGGGCCAAGGCACCACCTCTCACAGTCTCTCCTC	353
<i>Db</i>	310	GTCTTTGTACTCTGGGGCCAAGAACCACTCTCACAGTCTCTCCTC	353

RESULT 11  
AR117135  
LOCUS AR117135 720 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 10 from patent US 6140075.

```

REFERENCE 1 (bases 1 to 1797)
AUTHORS Fitzgerald,D., Chaudhary,V.Kumar., Pastan,I.Harry.,
TITLE Walzmann,T.Alexander. and Queen,C.L.
JOURNAL Recombinant antibody-toxin fusion protein
FEATURES
    source      1..1797
                /organism="unknown"
                /mol_type="unassigned DNA"

ORIGIN
Query Match      75.5%; Score 267.2; DB 6; Length 1797;
Best Local Similarity 88.1%; Pred.No. 3.1e-76;
Matches 303; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

Qy 10 CTGCTCGAGCAGCCTGGGGCTGAACCTGGGCAAAAACCTGGGGCCCTCAGTGAAGATGTCTCTGC 69
    |||
Db 10 CAGCTGCAGCAGTCTGGGGCTGAACCTGGGCAAAAACCTGGGGCCCTCAGTGAAGATGTCTCTGC 69

Qy 70 AAGCTTCTCGCTACACCTTTTACTACTACTGATTCACCTGGGTGAAACAGAGGCCCTGGA 129
    |||
Db 70 AAGCTTCTCGCTACACCTTTTACTAGCTACAGGATGCACCTGGGTAAACAGAGGCCCTGGA 129

Qy 130 CAGGGCTCGAAATGGATTTGGATACATTAATCTCCACACCTGGTTCCACTTCTTACAATCAG 189
    |||
Db 130 CAGGGCTCGAAATGGATTTGGATATATTAATCTCCACACTGGGTATCTGAATACAATCAG 189

Qy 190 GACTTTTCAGACAGGGCCACTTTTGACGGCAGACAGAGTCTCTCCACAGCAGCCTACATGCAG 249
    |||
Db 190 AAGTTCAAGACAGAAGCCCAATTGACTGTCAGACAAATCTCTCCACAGCAGCCTACATGCAG 249

Qy 250 CTGACACAGCCTGCATCTCTGAGGACTCTTCACTAGTCTTATTACTGTGCAAGAGAGGGGTACGAC 309
    |||
Db 250 CTGACAGCCTGCATCTTGGAGACTCTGAGTCTTATTACTGTGCAAGAGAGGGGG--GGG 306

Qy 310 GGGTTTGACTCTCGGGGCCAAGCAGACACTCTCACAGTCTCTCTC 353
    |||
Db 307 GTCTTTGACTACTCGGGGCCAAGGAAGCACTCTCACAGTCTCTCTC 350

RESULT 13
MUSIGHAAB
LOCUS Mouse Ig rearranged gamma-chain mRNA, clone AN05g.
DEFINITION M19898 J03832
ACCESSION M19898.1 GI:194554
VERSION C-region; immunoglobulin gamma-chain; immunoglobulin heavy chain;
KEYWORDS processed gene.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 49)
AUTHORS Rule,G.S.
JOURNAL Unpublished (1988)
REFERENCE 2 (bases 50 to 470)
AUTHORS Leahy,D.J., Rule,G.S., Whittaker,M.M. and McConnell,H.M.
TITLE Sequences of 12 monoclonal anti-dinitrophenyl spin-label antibodies
    for NMR studies
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (11), 3661-3665 (1988)
MEDLINE 88234486
PUBMED 3375235
COMMENT Original source text: Mus musculus (strain BALB/c, sub_species
    domesticus) cDNA to mRNA.
    Draft entry and computer-readable sequence [2] kindly submitted by
    G.Rule, 20-JUL-1988.

FEATURES
    source      1..470
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="BALB/c"
                /sub_species="domesticus"
                /db_xref="taxon:10090"

```

```
/cell_type="mature B cell"
50..>470
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
/cell_type="mature B cell"
55..>478
/notes="precursor"
/codon_start=1
/product="immunoglobulin gamma-chain"
/protein_id="AAA37956.1"
/db_xref="GI:553950"
/translation="MERHWIFLFLFSVTAGVHSQVPOQSGAEIAPGASVKMSCKAS
GYSTRYMHVWKORPGGLEWIGVINPTGYTVYNOKFKDKATLTADKSSSTAYMQL
SSLTSEDSAVYICAHYGRSYFDYWGQGTLLTVSSAKRP"
50..106
107..>470
/product="immunoglobulin gamma-chain"

sig_peptide
mat_peptide

ORIGIN
Query Match 75.3%; Score 266.6; DB 10; Length 470;
Best Local Similarity 85.8%; Pred. No. 4e-76;
Matches 296; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 10 CTGCTCGAGCAGCCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69
Db 113 CAGCCTCAGCAGTCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 172

Qy 70 AAGGCTTCTGGCTACACCTTTACTAACTACTGGATTCACTGGGTGAAACAGAGGCTGGA 129
Db 173 AAGGCTTCTGGCTACAGCTTTACCGCTACTGGATGCACTGGTGAACAGAGGCTGGA 232

Qy 130 CAGGCTTCTGAATGGATTGATATTAATCTTGCCACTGGTTCACCTTTTACAATCAG 189
Db 233 CAGGCTTCTGAATGGATTGATATTAATCTTAGCACTGGTTACTTGAATCAATCAG 292

Qy 190 GACTTTCAGGACAGGCGCACCTTTCAGCCGACAGATGCTCCACAGCCTACATGAG 249
Db 293 AAGTTCAGGACAGGCGCACCTTTCAGCCGACAGATGCTCCACAGCCTACATGAA 352

Qy 250 CTGACACGCTGACATCTGAGACTCTTTCAGTCTATTACTGTGCAAGAGAGGGTACGAC 309
Db 353 CTGACACGCTGACATCTGAGACTCTTTCAGTCTATTACTGTGCAAGAGAGGGTACG 412

Qy 310 GGGTTTGACTCTCTGGGGCCAGGACCACTCTCAGTCTCTCTCA 354
Db 413 TACTTTGACTCTCTGGGGCCAGGACCACTCTCAGTCTCTCTCA 457

RESULT 14
MUSIGHAAN 478 bp mRNA linear ROD 26-MAR-1994
LOCUS Mouse Ig rearranged gamma-chain mRNA, clone AN05g.
DEFINITION M19897 J03832
ACCESSION M19897.1 GI:194588
VERSION C-region; immunoglobulin gamma-chain; immunoglobulin heavy chain;
KEYWORDS processed gene.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Rule, G.S.
1 (bases 1 to 54)
Unpublished (1988)
2 (bases 55 to 478)
Leahy,D.J., Rule,G.S., Whittaker,M.M. and McConnell,H.M.
REFERENCE Sequences of 12 monoclonal anti-dinitrophenyl spin-label antibodies
AUTHORS for NMR studies
TITLE Proc. Natl. Acad. Sci. U.S.A. 85 (11), 3661-3665 (1988)
JOURNAL 88234486
MEDLINE 3375235
PUBMED Original source text: Mus musculus (strain BALB/c, sub_species
domesticus) cDNA to mRNA.
COMMENT Draft entry and computer-readable sequence [2] kindly submitted by
G.Rule, 20-JUL-1988.
FEATURES Location/Qualifiers
source 1..478
/organism="Mus musculus"
```

(E-mail:yabiko@mascat.nihon-u.ac.jp, Tel:+81-47-368-6111,  
Fax:+81-47-361-8880)

```
FEATURES
  source
    1..783
      /organism="Mus musculus"
      /mol_type="mRNA"
      /db_xref="taxon:10090"
  gene
    1..783
      /gene="MDABG2-4"
  CDS
    1..783
      /gene="MDABG2-4"
      /codon_start=1
      /product="single chain antibody ScFv"
      /protein_id="BAA22843.1"
      /db_xref="GI:2541868"
      /translation="MAQVQLQDSGAELKPGASVKMSCKASGYFTNYWHWIKQRPQ
      QGLEWIGYINPSTGYINQPKDKATLTADKSSSTAYMQLSSLTSDSAVYYCAKS
      GGFYFDYWGSGTIVTVSSGGSGGGGGGSDIELTQSPASLSVSGVETVITTCRA
      SENIYSLNLAHYQQKQSPQLLVYAATNLADGPSRFSGSGSGTQYSLKINSLSQSEDF
      GSYVCOHFWGTPTFGGGTKLELKAAGAPVPPDLEPRPA"
  V_region
    1..360
      /gene="MDABG2-4"
      /product="MAB heavy chain variable region"
  misc_feature
    361..405
      /gene="MDABG2-4"
      /note="linker peptide"
  V_region
    406..735
      /gene="MDABG2-4"
      /product="MAB light chain variable region"
  misc_feature
    736..774
      /gene="MDABG2-4"
      /note="E tag marker peptide"
```

## ORIGIN

```
Query Match          75.3%; Score 266.6; DB 10; Length 783;
Best Local Similarity 87.8%; Pred. No. 4.3e-76;
Matches 303; Conservative 0; Mismatches 39; Indels 3; Gaps 1;

Qy 13 CTGAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGCGCTCAGTGAAGATGTCCTGCAAG 72
Db 16 CTGAGCAGCTCTGGGGCTGAACCTGGCAAAACCTGGGCGCTCAGTGAAGATGTCCTGCAAG 75

Qy 73 GCTTCTGGCTACACCTTTTACTACTTACTGGATTCACTGGGTGAAACAGAGCGCTGGACAG 132
Db 76 GCTTCTGGCTACACCTTTTACTACTTACTGGATGCACTGGATATAAACAGAGCGCTGGACAG 135

Qy 133 GGTCTGAATGGATTGGATACATTAACTCTGCCACTGGTTCCACTTCTTACAATCAGGAC 192
Db 136 GGTCTGAATGGATTGGATACATTAACTCTGCCACTGGTTCCACTTCTTACAATCAGGAC 195

Qy 193 TTTCAAGACAGGGCCACTTTGACCGCAGACAAGTCTCCACACAGCCTACATGCACTG 252
Db 196 TTTCAAGACAGGGCCACTTTGACCTGACAGCAAAATCTCCAGCAGCGCTACATGCAACTG 255

Qy 253 ACCAGCCTGACATCTGAGGACTCTTTCAGTC---TATTACTGTGCAAGAGGGGTACGAC 309
Db 256 AGCAGCCTGACATCTGAGGACTCTGCACTCTATTATTACTGTGCAAAATCAGGGGGTGAC 315

Qy 310 GGGTTTGACTCTGGGGCCCAAGGACCACTCTCACAGTCTCCTCA 354
Db 316 TACTTTGACTACTGGGGCCCAAGGACCACTCTCACAGTCTCCTCA 360
```

Search completed: June 4, 2005, 08:29:06  
Job time : 3312.69 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 03:44:34 ; Search time 704.874 Seconds  
(without alignment)  
2972.996 Million cell updates/sec

Title: US-10-089-452-1

Perfect score: 354

Sequence: 1 gaggtcgagctgctcgagca.....ccactctcacagtctctctca 354

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_l6Dec04:\*

- 1: Geneseq\_l980s:\*
- 2: Geneseq\_l990s:\*
- 3: Geneseq\_l2000s:\*
- 4: Geneseq\_l2001as:\*
- 5: Geneseq\_l2001bs:\*
- 6: Geneseq\_l2002as:\*
- 7: Geneseq\_l2002bs:\*
- 8: Geneseq\_l2003as:\*
- 9: Geneseq\_l2003bs:\*
- 10: Geneseq\_l2003cs:\*
- 11: Geneseq\_l2003ds:\*
- 12: Geneseq\_l2004as:\*
- 13: Geneseq\_l2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	354	100.0	354	4	Aaf88151 H. pylori
2	354	100.0	354	4	Aaf88094 H. pylori
3	268.2	75.8	406	9	Abt43568 Murine an
4	268.2	75.8	443	1	Aan91820 DNA seque
5	267.2	75.5	719	3	Aaa60938 Anti-TAC
6	267.2	75.5	720	4	Aac60427 Anti-TAC
7	267.2	75.5	1797	2	Aav00687 Fusion ge
8	267.2	75.5	1797	3	Aaa10398 Recombina
9	266.8	75.4	348	10	Aad54525 atacvH DN
10	266.6	75.3	408	2	Aaq66846 Sequence
11	264.2	74.6	351	2	Aat39557 Variable
12	264.2	74.6	351	2	Aat69541 Anti-huma
13	264.2	74.6	354	2	Aax00882 Mouse der
14	261.2	73.8	898	2	Aav73335 Mouse OKT
15	261.2	73.8	1794	2	Aav73337 Mouse bis
16	261.2	73.8	6091	12	Adm32741 Nucleotid
17	260.2	73.5	351	2	Aaq39234 GMP-140 M
18	260.2	73.5	357	12	Adq91062 Murine OK
19	260.2	73.5	357	13	Adr42624 Mouse ant
20	260.2	73.5	417	12	Adq91060 Murine OK

21	260.2	73.5	1434	13	ADR42632	Adr42632 Mouse ant
22	260.2	73.5	1570	2	AAQ12637	Aaql2637 Monoclonal
23	260.2	73.5	1570	12	ADQ91058	Adq91058 Murine OK
24	260.2	73.5	1617	13	ADR42631	Adr42631 Mouse ant
25	260.2	73.5	1731	10	ADE52215	Ades52215 E3B1 codi
26	260.2	73.5	1982	13	ADR42629	Adr42629 Mouse ant
27	260.2	73.5	1982	13	ADR42628	Adr42628 Mouse ant
28	260.2	73.5	8078	10	ADE52213	Ades52213 pGIEN-EH3
29	259.6	73.3	1653	3	AAZ43432	Aaz43432 Fv-antibo
30	259.6	73.3	1698	3	AAZ43431	Aaz43431 Fv-antibo
31	259.6	73.3	1817	8	ACC79606	Acc79606 Plasmid p
32	259.6	73.3	1817	8	ACC79607	Acc79607 Plasmid p
33	259.2	73.2	1797	2	AAx00812	Aax00812 Vm-VI-lin
34	258	72.9	642	2	AAZ28805	Aaz28805 Coding se
35	257.2	72.7	1479	13	ADS99448	Adsg99448 Antibody
36	257	72.6	357	2	AAV54787	Aav54787 DNA encod
37	257	72.6	422	2	AAV54863	Aav54863 Murine an
38	257	72.6	422	2	AAV56413	Aav56413 Murine IC
39	257	72.6	422	2	AAV69185	Aav69185 Murine an
40	257	72.6	422	2	AAx21879	Aax21879 DNA encod
41	257	72.6	422	3	AAZ24327	Aaz24327 Murine an
42	257	72.6	422	3	AAA97155	Aaa97155 DNA encod
43	257	72.6	422	3	AAA08301	Aaa08301 Murine IC
44	257	72.6	422	6	ABK09344	Abk09344 cDNA enco
45	257	72.6	422	10	ADG25727	Adg25727 Mouse ant

#### ALIGNMENTS

##### RESULT 1

AAF88151  
ID AAF88151 standard; DNA; 354 BP.

AC AAF88151;

DT 17-JUL-2001 (first entry)

DE H. pylori catalase antibody HP25/2m/2H10 V region heavy chain DNA.

KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;

KW acid-resistant microorganism; complementarity determining region; CDR;

KW feces; heavy chain; light chain; ds.

XX Unidentified.

OS WO200127612-A2.

FN PD 19-APR-2001.

XX 12-OCT-2000; 2000MO-EP010057.

XX 12-OCT-1999; 99EP-00120351.

XX 16-MAR-2000; 2000EP-00105592.

XX 31-MAR-2000; 2000EP-00107028.

XX 10-MAY-2000; 2000EP-00110110.

PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

PI Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;

XX WPI; 2001-282086/29.

DR P-PSDB; AAB86105.

XX Detecting infections by acid-resistant microorganisms, particularly for  
diagnosing Helicobacter pylori, comprises immunochromatographic detection  
of antigen in feces.

PS Claim 32; Fig 1; 90pp; German.

XX This invention describes a novel method for detecting infection by an

acid-resistant microorganism (A), in a mammal, using

immunochromatography. The method is used to diagnose infection by an acid

CC -resistant microorganism (A), in a mammal, such as Helicobacter,  
CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
CC H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,  
CC inexpensive and non-invasive, and may indicate the stage of infection. A  
CC test strip used in the method may include a filter to eliminate particles  
CC present in the sample and only a single receptor provides a reasonably  
CC secure diagnosis, with specificity and selectivity improved by detecting  
CC several epitopes (of catalase) or different antigens (catalase and beta-  
CC urease). The method can be automated. This sequence encodes a  
CC Helicobacter pylori catalase derived antibody V-region heavy chain  
CC fragment used to illustrate the method of the invention

XX Sequence 354 BP; 85 A; 98 C; 92 G; 79 T; 0 U; 0 Other;

Query Match 100.0%; Score 354; DB 4; Length 354;  
Best Local Similarity 100.0%; Pred. No. 2.5e-96;  
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GAGGTGACGCTGCTCGAGCAGCCTGGGGCTGAATCTGCAAACTGGGGCTCAGTGAAG 60  
Db |||||  
1 GAGGTGACGCTGCTCGAGCAGCCTGGGGCTGAATCTGCAAACTGGGGCTCAGTGAAG 60  
Qy 61 ATGTCTCTGCAGGGCTTCTGGCTACACCTTTACTACTGATTCAGTGGTGAACAG 120  
Db |||||  
61 ATGTCTCTGCAGGGCTTCTGGCTACACCTTTACTACTGATTCAGTGGTGAACAG 120  
Qy 121 AGGCTGACAGGGCTTCTGAAATGGATGATACATTAATCTGCACCTGGTTCACCTTCT 180  
Db |||||  
121 AGGCTGACAGGGCTTCTGAAATGGATGATGATACATTAATCTGCACCTGGTTCACCTTCT 180  
Qy 181 TACAATCAGGACTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCCACACAGCC 240  
Db |||||  
181 TACAATCAGGACTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCCACACAGCC 240  
Qy 241 TACATGACGACTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCCACACAGCC 300  
Db |||||  
241 TACATGACGACTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCCACACAGCC 300  
Qy 301 GGGTACGACGGGTTTGACTCTCTGGGGCCAGGCAACCACTCTCACAGTCTCCTCA 354  
Db |||||  
301 GGGTACGACGGGTTTGACTCTCTGGGGCCAGGCAACCACTCTCACAGTCTCCTCA 354

RESULT 2  
AAF88094  
ID AAF88094 standard; DNA; 354 BP.  
XX  
AC AAF88094;  
XX  
DT 17-JUL-2001 (first entry)  
XX  
DE H. pylori catalase antibody HP25/2m 2H10 V-region heavy chain DNA.  
XX  
KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
KW antibacterial; complementarity determining region; V-region; db.  
XX  
OS Unidentified.  
XX  
PN WO200127613-A2.  
XX  
PD 19-APR-2001.  
XX  
PP 12-OCT-2000; 2000WO-EP010058.  
XX  
PR 12-OCT-1999; 99EP-00120351.  
PR 16-MAR-2000; 2000EP-00105592.  
PR 31-MAR-2000; 2000EP-00107028.  
PR 10-MAY-2000; 2000EP-00110110.  
XX  
PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
XX  
PI Reiter C, Cullmann G, Heppner P, Ringels A, Mueller H, Haindl E;

XX WPI; 2001-282087/29.  
DR P-PSDB; AAB86067.  
XX  
PT Detecting infections by acid-resistant microorganisms, particularly for  
PT diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
PT sample.  
XX  
PS Claim 28; Fig 1; 89pp; German.

XX This invention describes a novel method for detecting, in a mammal,  
CC infection by an acid-resistant microorganism (A) which comprises reacting  
CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
CC complex is formed with Ag, and the formation of a complex detected. R are  
CC specific for an Ag which, after passage through the intestines, at least  
CC in some mammals, retains a native (or corresponding) structure against  
CC which the mammal produces antibodies (when immunized or infected with  
CC (A), or its extracts, lysates or derived proteins (or fragments) or  
CC synthetic peptides). The products of the invention have antibacterial  
CC activity. The method is used to diagnose infection by Helicobacter,  
CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the  
CC progress of treatment. Receptors, particularly antibodies, directed  
CC against Ag can be used therapeutically for treatment of infections. The  
CC method requires only one R to provide a reasonably secure diagnosis  
CC (although use of two R improves sensitivity), so is relatively  
CC inexpensive and more easily standardized. Also it is direct, non-  
CC invasive, suitable for automation and may indicate the stage of an  
CC infection. This sequence encodes a Helicobacter pylori anti-catalase  
CC derived antibody HP25/2m 2H10 V-region heavy chain fragment which is  
CC described in the method of the invention

SQ Sequence 354 BP; 85 A; 98 C; 92 G; 79 T; 0 U; 0 Other;

Query Match 100.0%; Score 354; DB 4; Length 354;  
Best Local Similarity 100.0%; Pred. No. 2.5e-96;  
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGTGACGCTGCTCGAGCAGCCTGGGGCTGAATCTGCAAACTGGGGCTCAGTGAAG 60  
Db |||||  
1 GAGGTGACGCTGCTCGAGCAGCCTGGGGCTGAATCTGCAAACTGGGGCTCAGTGAAG 60  
Qy 61 ATGTCTCTGCAGGGCTTCTGGCTACACCTTTACTACTGATTCAGTGGTGAACAG 120  
Db |||||  
61 ATGTCTCTGCAGGGCTTCTGGCTACACCTTTACTACTGATTCAGTGGTGAACAG 120  
Qy 121 AGGCTGACAGGGCTTCTGAAATGGATGATGATACATTAATCTGCACCTGGTTCACCTTCT 180  
Db |||||  
121 AGGCTGACAGGGCTTCTGAAATGGATGATGATGATACATTAATCTGCACCTGGTTCACCTTCT 180  
Qy 181 TACAATCAGGACTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCCACACAGCC 240  
Db |||||  
181 TACAATCAGGACTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCCACACAGCC 240  
Qy 241 TACATGACGACTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCCACACAGAG 300  
Db |||||  
241 TACATGACGACTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCCACACAGAG 300  
Qy 301 GGGTACGACGGGTTTGACTCTCTGGGGCCAGGCAACCACTCTCACAGTCTCCTCA 354  
Db |||||  
301 GGGTACGACGGGTTTGACTCTCTGGGGCCAGGCAACCACTCTCACAGTCTCCTCA 354

RESULT 3  
ABT43568  
ID ABT43568 standard; cDNA; 406 BP.  
XX  
AC ABT43568;  
XX  
DT 09-OCT-2003 (first entry)  
XX  
DE Murine anti-CD25 heavy chain partial cDNA sequence.

KW Streptavidin fusion protein; SA: Streptomyces avidinii; biotin; scFvSA;  
 KW anti-CD25 antibody; antigen binding fragment; cytostatic; gene therapy;  
 KW cell-specific targeting agents; single chain; tumour; mouse; murine;  
 KW gene; ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..405  
 FT FT /\*tag= a  
 FT FT /partial  
 FT FT /product= "Murine anti-CD25 heavy chain"  
 FT FT /note= "No stop codon"  
 XX  
 FN WO2003050260-A2.  
 XX  
 XX 19-JUN-2003.  
 XX  
 XX 06-DEC-2002; 2002WO-US039429.  
 XX  
 PR 07-DEC-2001; 2001US-00013173.  
 PR 17-MAY-2002; 2002US-00150762.  
 PR 16-SEP-2002; 2002US-00244821.  
 XX  
 XX (NEOR-) NEORX CORP.  
 XX  
 XX Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;  
 PI Dearstyn EA;  
 XX  
 XX WPI; 2003-532908/50.  
 DR P-PSDB; ABJ39020.  
 XX  
 PT New vector construct for expressing genomic streptavidin fusion proteins  
 PT useful as diagnostic markers or as cell-specific targeting agents.  
 XX  
 XX Example 16; Page 153; 156pp; English.  
 XX  
 CC This invention relates to novel vector constructs for the expression of  
 CC streptavidin fusion proteins. Streptavidin (SA) is produced by  
 CC Streptomyces avidinii and specifically binds water-soluble biotin. The  
 CC vectors comprise a nucleic acid sequence encoding genomic streptavidin, a  
 CC promoter operatively linked to the nucleic acid sequence and a cloning  
 CC site for insertion of a second nucleic acid sequence encoding an anti-  
 CC CD25 antibody or its antigen binding fragment. The fusion proteins  
 CC encoded are known as single chain antibody-genomic streptavidin fusion  
 CC proteins (scFvSA). The vectors may have cytostatic activity when used in  
 CC gene therapy. The vectors may be useful in expressing genomic  
 CC streptavidin fusion cassettes. The fusion proteins may be used as  
 CC diagnostic markers or as cell-specific targeting agents. These may also  
 CC be used in treating tumours. The present sequence is the partial cDNA  
 CC sequence of the murine anti-CD25 heavy chain (VH) used during  
 CC construction of the anti-CD25 (anti-TAC) single chain antibody-genomic  
 CC streptavidin fusion in example 16 of the specification  
 XX  
 SQ Sequence 406 BP; 105 A; 100 C; 106 G; 95 T; 0 U; 0 Other;  
 Query Match 75.8%; Score 268.2; DB 9; Length 406;  
 Best Local Similarity 88.1%; Pred. No. 1.8e-70;  
 Matches 304; Conservative 0; Mismatches 38; Indels 3; Gaps 1;  
 10 CTGCTCAGCAGCTGGGGCTGAACCTGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
 64 CAGCTTCAGCAGTCTGGGGCTGGAACCTGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 123  
 70 AAGGCTTCTGGCTACACCTTTTACTTAACCTACTGGATTCTACTGGTGAACACAGAGCCTGGA 129  
 124 AAGGCTTCTGGCTACACCTTTTACTTAGCTACAGGATGCACTGGTGAACACAGAGCCTGGA 183  
 130 CAGGCTCTGAATGGATTGGATACATTAACTCTGCACCTGGTTCACCTTCTTACAATCAG 189  
 184 CAGGCTCTGAATGGATTGGATATTAATCTTAGCAGCTGGGTACTGTAATCAATCAG 243

190 GACTTTGAGGAGGGCCACTTTGACCGCAGCAAGTCTCTCCACACAGCCTACATGCAG 249  
 244 AAGTTCAAGGACAAGGCCACATTCAGTCGACACAATCTCCAGCAGCCTACATGCAA 303  
 250 CTGACCCAGCCTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGAGAGGGTACGAC 309  
 304 CTGAGCAGCCTGACATTTGAGGACTCTGCACTCTATTACTGTGCAAGAGGGGGG---GGG 360  
 310 GGGTTTGACTCTCTGGGGCCAAAGGCACCACTCTCTCAGTCTCTCTCA 354  
 361 GTCTTTGACTACTGGGGCCAAAGGAACCACTCTCTCAGTCTCTCTCA 405  
 RESULT 4  
 AAN91820  
 ID AAN91820 standard; DNA; 443 BP.  
 XX  
 AC AAN91820;  
 XX  
 DT 24-OCT-2003 (revised)  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 22-MAR-1990 (first entry)  
 XX  
 DE DNA sequence of the V and J regions of the anti-Tac heavy chain.  
 XX  
 KW Anti-Tac monoclonal antibody; heavy chain; V region; J region;  
 KW interleukin-2 receptor; human T-cells; chimeric antibody.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 38..442  
 FT FT /\*tag= a  
 XX  
 PN WO8909622-A.  
 XX  
 PD 19-OCT-1989.  
 XX  
 XX 13-APR-1989; 89WO-US001578.  
 PR 15-APR-1988; 88US-00181862.  
 XX  
 PA (PROT-) PROTEIN DES LABS IN.  
 XX  
 XX Queen CL;  
 WPI; 1989-324085/44.  
 P-PSDB; AAP93151.  
 XX  
 PT New interleukin-2 receptor-specific chimeric antibodies - useful for  
 PT treating T-cell mediated disorders.  
 XX  
 PS Disclosure; Fig 2; 46pp; English.  
 CC  
 CC The anti-Tac heavy chain is produced by ATCC CRL 9688. It is used in a  
 CC mouse/human chimeric antibody, which is specifically reactive with  
 CC interleukin-2 receptors on human T-cells. (Updated on 25-MAR-2003 to  
 CC correct PA field.) (Updated on 27-AUG-2003 to correct OS field.) (Updated  
 CC on 24-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 443 BP; 117 A; 114 C; 110 G; 102 T; 0 U; 0 Other;  
 Query Match 75.8%; Score 268.2; DB 1; Length 443;  
 Best Local Similarity 88.1%; Pred. No. 1.9e-70;  
 Matches 304; Conservative 0; Mismatches 38; Indels 3; Gaps 1;  
 10 CTGCTCAGCAGCCTGGGGCTGAACCTGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
 101 CAGCTTCAGCAGTCTGGGGCTGGAACCTGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 160  
 70 AAGGCTTCTGGCTACACCTTTTACTTAACCTACTGGATTCTACTGGTGAACACAGAGCCTGGA 129

Db 161 AAGGCTTCTGGCTACACCTTTACTAGCTACAGGATGCACCTGGGTAAACACAGAGGCGCTGGA 220  
Qy 130 CAGGGTCTGAATGGATTGGATACATTAATCCTGCCACTGGTTCCACTTCTTACAATCAG 189  
Db 221 CAGGGTCTGGATGGATTGGATATATTATCTTAGCACTGGGTATATCTGATACAATCAG 280  
Qy 190 GACTTTTCAGGACAGGGGCACCTTTGACCGCGAGACAAGTCTCCACACAGCGCTACATGAG 249  
Db 281 AAGTTCAAGGACAAAGGCCACATTTGACTGCGAGACAAATCTCCAGCACAGCGCTACATGCAA 340  
Qy 250 CTGACACGCGCTGACATCTGAGGACTCTTCACTTACTTACTGTGCAAGAGGGGTACGAC 309  
Db 341 CTGAGCAGCGCTGACATTTGAGGACTCTGCGAGTCTATTACTGTGCAAGAGGGGGG---GGG 397  
Qy 310 GGGTTTGACTCTCTGGGCGCCAAAGGCCACCACTCTCACAGTCTCCTCA 354  
Db 398 GTCTTTGACTACTGGGCGCCAAAGGACCACCTCTCACAGTCTCCTCA 442

RESULT 5  
AAA60938  
ID AAA60938 standard; DNA; 719 BP.  
XX  
AC AAA60938;  
XX  
XX 08-NOV-2000 (first entry)  
XX  
XX Anti-TAC sfv DNA sequence SEQ ID NO:7.  
XX  
XX Tobacco; extensin; signal; ext26; ext21; PCR; immunological; antibody;  
XX mammalian; ds.  
XX  
XX Homo sapiens.  
XX  
XX US6080560-A.  
XX  
XX 27-JUN-2000.  
XX  
XX 25-JUL-1994; 94US-00279772.  
XX  
XX 25-JUL-1994; 94US-00279772.  
XX  
XX (MONS ) MONSANTO CO.  
XX  
XX Russell DR, Fuller JT;  
XX  
XX WPI: 2000-451206/39.  
XX P-PSDB; AAB12563.  
XX  
XX Producing mammalian antibody in plant cells involves transforming tobacco  
XX cells with genetic construct comprising a sequence encoding single chain  
XX antibody and isolating antibody from transformed cells.  
XX  
XX Example; Col 15-18; 11pp; English.

XX The present invention describes a method for obtaining a conformationally  
XX active mammalian antibody from plant cells in a culture involving  
XX transforming tobacco suspension cells with a DNA construct comprising a  
XX DNA sequence encoding a secretable mammalian single chain antibody,  
XX culturing the transgenic cells and isolating the accumulated single chain  
XX mammalian antibody from the tobacco cells. The method can be used for  
XX obtaining purified preparations of immunologically active  
XX conformationally intact mammalian antibody. High yield of commercially  
XX useful quantities of immunologically active mammalian antibody can be  
XX recovered by this method. The present sequence encodes an anti-TAC sfv  
XX sequence, which is used in the exemplification of the present invention

XX  
XX Sequence 719 BP; 174 A; 193 C; 190 G; 162 T; 0 U; 0 Other;  
Query Match 75.5%; Score 267.2; DB 3; Length 719;  
Best Local Similarity 88.1%; Pred. No. 4.4e-70;  
Matches 303; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

Qy 10 CTGCTCAGCAGCGCTGGGCTGAACTGGCAAAACCTGGGGCCTCAGTGAAGATCTCCTGC 69  
Db 13 CAGCTTCAGCAGTCTGGGCTGAACTGGCAAAACCTGGGGCCTCAGTGAAGATCTCCTGC 72  
Qy 70 AAGGCTTCTGGCTACACCTTTTACTTACTTACTTGGGTTGAAACACAGAGGCGCTGGA 129  
Db 73 AAGGCTTCTGGCTACACCTTTTACTTACTTACTTGGGTTGAAACACAGAGGCGCTGGA 132  
Qy 130 CAGGCTCTGAATGGATTGGATACATTAATCCTGCCACTGGTTCCACTTCTTACAATCAG 189  
Db 133 CAGGCTCTGAATGGATTGGATATATTATCTTAGCACTGGGTATATCTGATACAATCAG 192  
Qy 190 GACTTTTCAGCAGCGGCGCACTTTGACCGCGAGACAAGTCTCCACACAGCGCTACATGAG 249  
Db 193 AAGTTCAAGGACAAAGGCCACATTTGACTGCGAGACAAATCTCCAGCACAGCGCTACATGCAA 252  
Qy 250 CTGACACGCGCTGACATCTGAGGACTCTTCACTTACTTACTTGTGCAAGAGGGGTACGAC 309  
Db 253 CTGAGCAGCGCTGACATTTGAGGACTCTGCGAGTCTATTACTGTGCAAGAGGGGG---GGG 309  
Qy 310 GGGTTTGACTCTCTGGGCGCCAAAGGCCACCACTCTCACAGTCTCCTC 353  
Db 310 GTCTTTGACTACTGGGCGCCAAAGGACCACCTCTCACAGTCTCCTC 353

RESULT 6  
AAC60427  
ID AAC60427 standard; DNA; 720 BP.  
XX  
AC AAC60427;  
XX  
XX 19-FEB-2001 (first entry)  
XX  
XX Anti-TAC sfv coding sequence.  
XX  
XX Plant toxin; cancer; high yield; ds.  
XX  
XX Unidentified.  
XX  
XX US6140075-A.  
XX  
XX 31-OCT-2000.  
XX  
XX 29-JUL-1997; 97US-00902486.  
XX  
XX 25-JUL-1994; 94US-00279772.  
XX  
XX (MONS ) MONSANTO CO.  
XX  
XX Miller MJ, Russell DR, Fuller JT;  
XX WPI: 2001-040301/05.  
XX P-PSDB; AAB36826.  
XX  
XX Obtaining protein toxins from tobacco cells useful as therapeutic agent,  
XX by transforming a cell suspension with a DNA construct comprising a  
XX signal sequence for the secretion of the proteins.  
XX  
XX Example; Col 21-24; 19pp; English.

XX The present invention relates to obtaining secretable plant protein  
XX toxins from plant cells. The invention is useful for obtaining  
XX commercially useful and higher quantities of plant protein toxins. These  
XX proteins may be used as a therapeutic agent or cytotoxin directed against  
XX specific cancer cells by covalently linking the toxin to an antibody that  
XX recognizes a particular antigenic determinant associated with the cancer  
XX cell. The new method allows production of relatively large quantities of  
XX polypeptide toxins in a heterologous expression system, where the  
XX specific yield and concentration of the desired protein are markedly  
XX higher than those reported in previous methods

XX  
XX Sequence 720 BP; 174 A; 193 C; 191 G; 162 T; 0 U; 0 Other;



Query Match 75.5%; Score 267.2; DB 4; Length 720;  
Best Local Similarity 88.1%; Pred. No. 4.e-70;  
Matches 303; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

QY 10 CTGCTCGAGCAGCCCTGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
DB 13 CAGCTTCAGCAGCTCTGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 72

QY 70 AAGGCTTCGGCTACACCTTTACTACTGGAATTCATCTGGTGAACAGAGGCTTGA 129  
DB 73 AAGGCTTCGGCTACACCTTTACTAGCTACAGGATGCACTGGTGAACAGAGGCTTGA 132

QY 130 CAGGCTCGAATGATGGATGATACATTAACTCTGCCACTGCTTCCACTTCTTACAATCAG 189  
DB 133 CAGGCTCGAATGATGGATGATATTAATCTAGCACTGGGTATCTGAATACATCAG 192

QY 190 GACTTTCAGGACAGGGCCACTTTGACCGCAGACAAAGTCTCCACACAGCCTCATGTCAG 249  
DB 193 AAGTTCAAGGACAAAGGCCACATTGACTGCAGACAAATCTCCAGCAGCCTACATGCA 252

QY 250 CTGACCGCTGACATCTGAGGACTCTTCACTCTTATTAATCTGTGCAAGAGGGGTACGAC 309  
DB 253 CTGACCGCTGACATCTGAGGACTCTGCACTCTTATTAATCTGTGCAAGAGGGGGG---GGG 309

QY 310 GGGTTTCACTCTGGGGCCAAAGGACCACTCTCACAGTCTCTC 353  
DB 310 GTCTTTGACTACTGGGGCCAAAGGACCACTCTCACAGTCTCTC 353

## RESULT 7

AAV00687  
ID AAV00687 standard; DNA; 1797 BP.

AC AAV00687;

17-OCT-2003 (revised)  
25-MAR-2003 (revised)  
27-MAR-1998 (first entry)

Fusion gene sequence encoding anti-Tac(Fv)-PE40 antibody fusion protein.

Antibody-toxin fusion protein; Pseudomonas exotoxin; PE;  
protein synthesis inhibition; interleukin-2 receptor; immunotoxin;  
antibody; autoimmune disease; cancer cell; ss.

Pseudomonas.  
Mammalia.  
Chimeric.

Key Location/Qualifiers  
CDS 1..1797

misc\_feature /\*tag= a  
1..352

misc\_feature /\*tag= b  
/note= "encodes mature variable domain of the heavy chain"

misc\_feature 353..396  
/\*tag= c  
/note= "linker"

misc\_feature 397..714  
/\*tag= d  
/note= "encodes mature variable domain of the light chain"

misc\_feature 715..1797  
/\*tag= e  
/note= "encodes amino acids 253-613 of PE40"

US5696237-A.

09-DEC-1997.

05-JUN-1995; 95US-00463163.

XX

PR 24-SEP-1986; 86US-00911227.  
PR 21-APR-1989; 89US-003411361.  
PR 08-APR-1992; 92US-00865722.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA (PROT-) PROTEIN DESIGN LAB INC.

PI Waldmann TA, Chaudhary VK, Pastan IH, Queen CL, Fitzgerald D;

XX WPI; 1998-041352/04.

DR P-PSDB; RAN37132.

XX Antibody toxin fusion proteins - useful as immuno:toxin for treating,  
e.g. lymphocytes associated with auto:immune disease.

PS Disclosure; Col 13-18; 20pp; English.

XX The present sequence encodes a protein that is an example of a novel  
antibody-Pseudomonas exotoxin (PE) fragment recombinant fusion protein.  
CC It is designated anti-Tac(Fv)-PE40. The present fusion gene encodes  
CC various domains of the heavy chain (the first 116 amino acids), a 15  
CC amino acid linker, VL-variable domain of the light chain (first 106 amino  
CC acids), and amino acids 253-613 of PE. This PE fragment possesses  
CC translocating and ADP-ribosylating activity. The fusion gene is cloned  
CC under the control of a T7 promoter linked to Shine-Delgarno region and  
CC initiation codon. The fusion protein selectively inhibits protein  
CC synthesis in cells expressing the interleukin-2 receptor. The antibody-  
CC PE40 fusion protein is used as an immunotoxin for killing cells bearing a  
CC receptor or antigen to which the antibody binds, e.g. lymphocytes  
CC associated with autoimmune diseases or cancer cells. (Updated on 25-MAR-  
CC 2003 to correct PA field.) (Updated on 17-OCT-2003 to standardise OS  
CC field)

XX SQ Sequence 1797 BP; 336 A; 598 C; 571 G; 292 T; 0 U; 0 Other;

Query Match 75.5%; Score 267.2; DB 2; Length 1797;  
Best Local Similarity 88.1%; Pred. No. 6e-70;  
Matches 303; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

QY 10 CTGCTCGAGCAGCCTGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
DB 10 CAGCTGCAGCAGCTCTGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69

QY 70 AAGGCTTCGGCTACACCTTTACTACTGGAATTCATCTGGTGAACAGAGGCTTGA 129

DB 70 AAGGCTTCGGCTACACCTTTACTAGCTACAGGATGCACTGGGTGAACAGAGGCTTGA 129

QY 130 CAGGCTCTGAAATGGATTGGATACATTAATCTGCCACTGGTTCACCTTCTTACAATCAG 189

DB 130 CAGGCTCTGAAATGGATTGGATATTAATCTAGCACTGGGTATCTGAATACATCAG 189

QY 190 GACTTTTCAGGACAGGGCCACTTTTGACCGCAGACAAAGTCTCCACACAGCCTACATGCA 249

DB 190 AAGTTCAAGGACAGGGCCACATTCAGCTGCAGACAAATCTCCACACAGCCTACATGCA 249

QY 250 CTGACCGCTGACATCTGAGGACTCTTCACTCTTATTAATCTGTGCAAGAGGGGTACGAC 309

DB 250 CTGACCGCTGACATCTGAGGACTCTGCACTCTTATTAATCTGTGCAAGAGGGGGG---GGG 306

QY 310 GGGTTTCACTCTGGGGCCAAAGGACCACTCTCACAGTCTCTC 353

DB 307 GTCTTTGACTACTGGGGCCAAAGGACCACTCTCACAGTCTCTC 350

## RESULT 8

AAAI0398  
ID AAAI0398 standard; DNA; 1797 BP.

XX AAAI0398;

XX 15-SEP-2003 (revised)

DT 03-JUL-2000 (first entry)

XX

DE	Recombinant DNA encoding anti-Tac(Fv)-PE40 fusion protein.	
XX	Antibody-toxin fusion protein; single chain antibody; immunotoxin;	
KW	Pseudomonas exotoxin; anti-Tac(Fv)-PE40; targetted delivery;	
KW	interleukin-2 receptor; IL-2; helper T-lymphocyte; autoimmune disease;	
KW	leukaemia; ds.	
XX		
OS	Pseudomonas sp.	
OS	Mus sp.	
OS	Chimeric.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	1..1797
FT		/*tag= a
FT		/product= "Anti-Tac(Fv)-PE40 immunotoxin"
FT		/note= "No stop codon given in the specification"
FT		/transl_except= (pos:1174..1176, aa:Asp)
XX		
FN	US6051405-A.	
XX		
PD	18-APR-2000.	
XX		
PF	08-APR-1992;	92US-00865722.
XX		
PR	24-SEP-1986;	86US-00911227.
PR	21-APR-1989;	89US-00341361.
XX		
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.	
PA	(PROT-) PROTEIN DESIGN LABS INC.	
XX		
PI	Fitzgerald D, Chaudhary VK, Pastan IH, Queen CL, Waldmann TA;	
XX		
DR	WPI; 2000-363771/31.	
DR	P-PSDB; AAY87477.	
XX		
XX	Construct encoding recombinant scFV-toxin fusion protein to selectively	
FT	kill cells bearing antigens or receptors comprises DNA segment.	
XX		
PS	Claim 9; Col 9-12; 14pp; English.	
XX		
CC	This sequence represents DNA encoding a recombinant single chain antibody	
CC	fusion protein anti-Tac(Fv)-PE40, which comprises the heavy and light	
CC	chain variable regions (VH and VL) of an anti-Tac antibody and residues	
CC	253-613 of Pseudomonas exotoxin (PE). The anti-Tac single chain antibody	
CC	component of the immunotoxin binds to the p55 subunit (Tac antigen) of	
CC	the interleukin-2 (IL-2) receptor, which is present in large amounts on	
CC	helper T-lymphocytes. This enables the cytotoxic action of PE40 to be	
CC	targetted to these and other cell types which express the IL-2 receptor.	
CC	The recombinant immunotoxin of the invention may be used to treat a	
CC	variety of autoimmune diseases, including graft-versus-host disease,	
CC	organ transplant rejection, type I diabetes, multiple sclerosis,	
CC	rheumatoid arthritis, systemic lupus erythematosus and myasthenia gravis.	
CC	It may also be used in the treatment of leukaemia, and may be used in	
CC	vivo for the elimination of harmful cells from bone marrow before	
CC	transplant. Anti-Tac(Fv)-PE40 is more active on a molar basis than anti-	
CC	Tac antibody chemically conjugated to full-length PE (anti-Tac-PE) or to	
CC	PE40 (anti-Tac-PE40). (Updated on 15-SEP-2003 to standardise OS field)	
XX		
SQ	Sequence 1797 BP; 336 A; 599 C; 570 G; 292 T; 0 U; 0 Other;	
	Query Match	75.5%; Score 267.2; DB 3; Length 1797;
	Best Local Similarity	88.1%; Pred. No. 6e-70;
	Matches 303; Conservative 0; Mismatches 38; Indels 3; Gaps 1;	
QY	10 CTGCTCAGCAGCTGGGGCTGAACCTGGCAAACTGGGGCTCAGTGAAGATGTCCTGC	69
DB	10 CAGCTGCAGCAGCTGGGGCTGAACCTGGGGCTCAGTGAAGATGTCCTGC	69
QY	70 AAGGCTTCTGGCTACACCTTTACTGAATTCACCTGGGTGAACAGAGGCTTGG	129
DB	70 AAGGCTTCTGGCTACACCTTTACTAGTACAGGATGCATGGGTAAACAGAGGCTTGA	129
QY	130 CAGGGTCTGAATGGATTGGATACATTAATCCTGCCACTGGTTCCATCTTTACAATCAG	189

```
SQ Sequence 348 BP; 94 A; 85 C; 93 G; 76 T; 0 U; 0 Other;
Query Match 75.4%; Score 266.8; DB 10; Length 348;
Best Local Similarity 88.3%; Pred. No. 4.5e-70;
Matches 302; Conservative 0; Mismatches 37; Indels 3; Gaps 1;
QY 13 CTCGAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTCAAGATGTCCTGCAAG 72
DB 10 CTGACGAGCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTCAAGATGTCCTGCAAG 69
QY 73 GCTTCTGCTACACCTTTTACTAACTACTGGATTCTAGTGGTGAACAGAGGCGCTGGACAG 132
DB 70 GCTTCTGCTACACCTTTTACTAGCTACAGGATGCACTGGTGAACAGAGGCGCTGGACAG 129
QY 133 GGTCTGAATGGATTGGATACATTAATCTGCACTGGTTCACCTTCTTCAATCAGGAC 192
DB 130 GGTCTGAATGGATTGGATATATTAATCTTACGACTGGGTATCTGAATACAATCAGGAG 189
QY 193 TTTCAAGGACAGGGCCACTTTGACCGCAGACCAAGTCTCCACACAGCCTACATGAGCTG 252
DB 190 TTTCAAGGACAGGGCCACTTGAATGCTGACAGCAATCTCCAGCAGCCTACATGCACTG 249
QY 253 ACCAGCCTGACATCTGAGGACTCTTCACTTATTAATCTGCAAGAGAGGGGTACGACGG 312
DB 250 AGCAGCCTGACATTTGAGGACTCTGCACTTATTAATCTGCAAGAGAGGGGTACGACGG 306
QY 313 TTTGACTCTGGGGCCAGGACCACTCTCAGTCTCTCTCA 354
DB 307 TTTGACTCTGGGGCCAGGAACTCTCAGTCTCTCTCA 348

RESULT 10
AAQ66846
ID AAQ66846 standard; DNA; 408 BP.
AC AAQ66846;
XX
XX
XX 25-MAR-2003 (revised)
XX 31-DEC-1994 (first entry)
XX
XX Sequence encoding the heavy chain variable region of the mouse NA-8
XX antibody, including the signal sequence and mature chain.
XX
XX Antibody NA-8; human CD18; complementarity determining region; CDR; ss.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 1..408
XX /*tag= a
XX
XX WO9412214-A1.
XX
XX 09-JUN-1994.
XX
XX 30-NOV-1993; 93WO-US011611.
XX
XX 01-DEC-1992; 92US-00983949.
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
XX
XX Co MS, Landolfi NP;
XX
XX WPI; 1994-199973/24.
XX
XX P-PSDB; AAR59510.
XX
XX New humanised antibodies specific for CD18 - derived from new murine
XX antibody NA-8, prevent binding of neutrophils to endothelial cells,
XX useful for treating inflammation.
XX
XX Disclosure; Fig 1B; 50pp; English.
XX
XX The mouse antibody NA-8 binds to human CD18. cDNAs for the heavy chain
CC and light chain variable domain genes of NA-8 were cloned using anchored
CC PCR. The cDNA variable domain sequences and the deduced AA sequences are
CC shown in AAQ66845/R59509 and AAQ66846/R59510. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
XX SQ Sequence 408 BP; 104 A; 106 C; 98 G; 100 T; 0 U; 0 Other;
Query Match 75.3%; Score 266.6; DB 2; Length 408;
Best Local Similarity 85.8%; Pred. No. 5.5e-79;
Matches 296; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 10 CTGCTCGAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTCAAGATGTCCTGTC 69
DB 64 CAGCTTCAGCAGTCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTCAAGATGTCCTGTC 123
QY 70 AAGGCTTCTGCTACACCTTTTACTAACTACTGGATTCTAGTGGTGAACAGAGGCGCTGGA 129
DB 124 AAGGCTTCTGCTACACCTTTTACTAGTGGTGAACAGAGGCGCTGGA 183
QY 130 CAGGCTCTGAATGGATTGGATACATTAATCTGCACTGGTTCACCTTCTTACATCAG 189
DB 184 CAGGCTCTGGAATGGATTGGAAACATTAATCTTACGACTGCTTATACCGACTTACATCAG 243
QY 190 GACTTTCAAGCAGAGGGCCACTTTGACCGCAGACCAAGTCTCCACACAGCCTTACATGCA 249
DB 244 AACITTAACGACAGGGCCACTTGAATGCTGACAGCAGATCTCCACACAGCCTTACATGCA 303
QY 250 CTGACGAGCCTGACATCTGAGGACTCTTCACTTATTAATCTGCAAGAGAGGGGTACGAC 309
DB 304 CTGAGCAGCCTGACATCTGAGGACTCTGCACTTATTAATCTGCAAGAGGGGTGATTTC 363
QY 310 GGGTTTGACTCTGGGGCCAGGACCACTCTCAGTCTCTCTCA 354
DB 364 CTTATGGACTACTGGGGTCAAGGAACTCAGTCAACCGTCTCTCTCA 408

RESULT 11
AAT39557
ID AAT39557 standard; cDNA to mRNA; 351 BP.
XX
XX AAT39557;
XX
XX 19-MAY-1997 (first entry)
XX
XX Variable heavy chain cDNA for anti-human Fas ligand antibody NOK-5.
XX
XX Variable region; heavy chain; human; Fas ligand; monoclonal; antibody;
XX NOK-5 hybridoma; inhibition; apoptosis; assay; diagnosis; disease;
XX hepatitis; infectious mononucleosis; systemic lupus erythematosus; ds.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 1..351
XX /*tag= a
XX /note= "partial open reading frame"
XX
XX WO9629350-A1.
XX
XX 26-SEP-1996.
XX
XX 21-MAR-1996; 96WO-JP000734.
XX
XX 20-MAR-1995; 95JP-00087420.
XX
XX 27-OCT-1995; 95JP-00303492.
XX
XX (SUME ) SUMITOMO ELECTRIC IND CO.
XX
XX Kayagaki N, Yagita H, Okumura K, Nakata M;
XX
XX WPI; 1996-443140/44.
XX
XX P-PSDB; AAW00831.
XX
```

PT Monoclonal antibody specifically recognising the Fas ligand - useful for  
 XX the detection of Fas ligands either on cell surface or in solution.

PS Claim 27; Page 90; 133pp; Japanese.

XX The present sequence encodes the heavy chain variable region of the anti-  
 CC human Fas ligand monoclonal antibody (Mab) NOK-5. NOK-5 is produced by  
 CC the hybridoma NOK-5 (FERM BP-5048), which was prepared by immunising mice  
 CC with transformed human Fas ligand expressing COS cells, and fusing spleen  
 CC cells isolated from the mice with myeloma P3x63Ag8.653 (ATCC CRL-1580)  
 CC cells. The Mab recognises the human Fas ligand on the cell surface or in  
 CC solution. The Mab can be used to inhibit the apoptosis inducing cell surface  
 CC Fas ligand/Fas reaction. The Mab can also be used for a Fas ligand assay  
 CC in biological samples (e.g. human blood), especially for disease  
 CC diagnosis, e.g. hepatitis, infectious mononucleosis and systemic lupus  
 CC erythematosus

XX Sequence 351 BP; 96 A; 87 C; 90 G; 78 T; 0 U; 0 Other;

Query Match 74.6%; Score 264.2; DB 2; Length 351;  
 Best Local Similarity 86.8%; Pred. No. 2.7e-69;  
 Matches 303; Conservative 0; Mismatches 43; Indels 3; Gaps 1;

QY 9 GCTGCTCGAGCAGCTGGGCTGAACTGGGCAAAACCTGGGGCCTCAGTGAAGATGTCCTG 68  
 DB |||||  
 QY 3 GCAGCTGCAGGAGTCTGGGGCTGAACCGGCAAAACCTGGGGCCTCAGTGAAGATGTCCTG 62  
 DB |||||  
 QY 69 CAAGGCTTCTGGCTACACTTTACTTAATCTACTGATTCACCTGGTGAACACAGAGGCTGG 128  
 DB |||||  
 QY 63 CAAGGCTTCTGGCTACACTTTACTTAATCTACTGATTCACCTGGTGAACACAGAGGCTGG 122  
 DB |||||  
 QY 129 ACAGGGTCTGAAATGGATGGATACATTAATCTCTGCCACTGGTTCCTTCAATCAATCA 188  
 DB |||||  
 QY 123 ACAGGGTCTGAAATGGATGGATACATTAATCTCTGCCACTGGTTCCTTCAATCAATCA 182  
 DB |||||  
 QY 189 GGACTTTTCAGACAGGGCCACTTTGACCGGACAGCAAGTCTCTCCACACAGCCTACATGCA 248  
 DB |||||  
 QY 183 GAAGTTCAAGGACCAAGGCCACATTTGACTGCGACAGCAAAATCTCTCCAGCAGCCTACATGCA 242  
 DB |||||  
 QY 249 GCTGACCAAGCTGACATCTGAGGACTCTTCACTGATTCAGTGTCTGAGAGAGGG---GTA 305  
 DB |||||  
 QY 243 ACTAATCAGCCTGACATCTGAGGACTCTGCACTGATTTACTGTGCAAGAGGGGTAATTA 302  
 DB |||||  
 QY 306 CGACGGGTTTGACTCTCTGGGGCCAAAGCCACCACTCTCAGTCTCTCTCA 354  
 DB |||||  
 QY 303 CTACTACTTTGACTCTGGGGCCAAAGGACCAAGGCTCAGCGTCTCTCTCA 351  
 DB |||||

RESULT 12

AAAT69541  
 ID AAAT69541 standard; cDNA to mRNA; 351 BP.

XX AC AAAT69541;

XX DT 14-JAN-1998 (first entry)

XX DE Anti-human FasL antibody (NOK5) heavy chain variable region cDNA.

XX KW Heavy chain; variable region; mouse; murine; human; Fas ligand; FasL;  
 KW monoclonal antibody; Mab; hybridoma; treatment; hepatitis;  
 KW hepatitis B virus; HBV; hepatitis C virus; HCV; apoptosis; liver cell;  
 KW glutamate oxaloacetate; pyruvate transaminase; ds.

XX OS Mus sp.

XX FH Key Location/Qualifiers  
 FT mat\_peptide 1..351  
 FT /\*tag= a

XX PN WO9715326-A1.

XX PD 01-MAY-1997.

XX

PF 24-OCT-1996; 96WO-JP003089.

XX PR 27-OCT-1995; 95JP-00303491.

XX PA (SUME ) SUMITOMO ELECTRIC IND CO.

XX FI Seino K, Kayagaki N, Yagita H, Okumura K, Nakata M;

XX WPI; 1997-258767/23.

XX DR P-PSDB; AAW19017.

XX Anti-human Fas ligand antibody to treat hepatitis - controls apoptosis in  
 PT liver cells and improves liver function.

XX Claim 6; Page 40; 51pp; Japanese.

XX The present sequence encodes the heavy chain variable region of the  
 CC murine anti-human Fas ligand (FasL) monoclonal antibody (Mab) NOK5, which  
 CC is expressed by the hybridoma NOK5 (FERM BP-5044). The Mab can be used in  
 CC the preparation of a composition for the effective oral or parenteral  
 CC treatment of hepatitis, including hepatitis caused by hepatitis B or C  
 CC virus. The composition controls apoptosis in liver cells caused by the  
 CC binding of FasL to Fas expressing liver cells, and improves liver  
 CC function by improving blood glutamate oxaloacetate and pyruvate  
 CC transaminase levels. The composition is given in a dosage of 0.0001-1000,  
 CC preferably 0.01-600 mg/day. Spleen cells from mice immunised with FasL  
 CC expressing COS cells were fused with mouse myeloma cells to produce  
 CC hybridomas. The hybridomas were screened for anti-FasL activity, and the  
 CC active clones NOK1-5 isolated

XX Sequence 351 BP; 96 A; 87 C; 90 G; 78 T; 0 U; 0 Other;

Query Match 74.6%; Score 264.2; DB 2; Length 351;  
 Best Local Similarity 86.8%; Pred. No. 2.7e-69;  
 Matches 303; Conservative 0; Mismatches 43; Indels 3; Gaps 1;

QY 9 GCTGCTCGAGCAGCTGGGCTGAACTGGGCAAAACCTGGGGCCTCAGTGAAGATGTCCTG 68

DB |||||  
 QY 3 GCAGCTGCAGGAGTCTGGGGCTGAACCGGCAAAACCTGGGGCCTCAGTGAAGATGTCCTG 62

QY 69 CAAGGCTTCTGGCTACACTTTACTTAATCTACTGATTCAGTGTGAACACAGAGGCTGG 128

DB |||||  
 QY 63 CAAGGCTTCTGGCTACACTTTACTTAATCTACTGATTCAGTGTGAACACAGAGGCTGG 122

QY 129 ACAGGGTCTGAAATGGATGGATACATTAATCTCTGCCACTGGTTCCTTCAATCAATCA 188

DB |||||  
 QY 123 ACAGGGTCTGAAATGGATGGATACATTAATCTCTGCCACTGGTTCCTTCAATCAATCA 182

QY 189 GGACTTTTCAGACAGGGCCACTTTGACCGGACAGCAAGTCTCTCCACACAGCCTACATGCA 248

DB |||||  
 QY 183 GAAGTTCAAGGACCAAGGCCACATTTGACTGCGACAGCAAAATCTCTCCAGCAGCCTACATGCA 242

QY 249 GCTGACCAAGCTGACATCTGAGGACTCTTCACTGATTCAGTGTCTGAGAGAGGG---GTA 305

DB |||||  
 QY 243 ACTAATCAGCCTGACATCTGAGGACTCTGCACTGATTTACTGTGCAAGAGGGGTAATTA 302

QY 306 CGACGGGTTTGACTCTCTGGGGCCAAAGCCACCACTCTCAGTCTCTCTCA 354

DB |||||  
 QY 303 CTACTACTTTGACTCTGGGGCCAAAGGACCAAGGCTCAGCGTCTCTCTCA 351

XX RESULT 13

XX AAAX00882

XX ID AAAX00882 standard; DNA; 354 BP.

XX AC AAAX00882;

XX DT 29-MAR-1999 (first entry)

XX DE Mouse derived RT3 phase antibody heavy chain pattern D genetic sequence.

XX KW Catalytic; antibody; phage display; immunising; phage expression vector;

XX produg; scFV; ss.



Db 219 ACAGGGTCTGAATGGATGGATACATTAATCCTAGCGGTGGTTATATACTAATCAATCA 278  
Qy 189 GGACTTTTCAGGACAGGCGCACTTTGACCGCAGACAAAGTCCTCCACACAGCCTACATGCA 248  
Db 279 GAAGTTCAAGGACAGGCGCACTTACGATACAGACAAATCCTCCAGCAGCCTACATGCA 338  
Qy 249 GCTGACCAAGCCTGACATCTGAGGACTCTTCAGTCTTATTACTGTGCAAGA-----GAGGG 302  
Db 339 ACTGAGCAGCCTGACATCTGAGGACTCTGAGTCTTATTACTGTGCAAGATATTATGATGA 398  
Qy 303 GTAGACGGGTTTGACTCTCGGGCCAGGACCACTCTACAGTCTCCTCA 354  
Db 399 TCATTACAGCCTTGACTACTCGGGCCAGGACCACTCTCAGTCTCCTCA 450

RESULT 15  
AAV73337  
ID AAV73337 standard; DNA; 1794 BP.  
XX AC AAV73337;  
XX AC AAV73337;  
XX 26-FEB-1999 (first entry)  
XX Mouse bispecific antibody construct variant OKT3/anti-CD19 DNA.  
XX OKT3; monoclonal antibody; MAb; point mutation; transplant rejection;  
XX organ recipient; diagnosis; tumour; therapy; diabody; anti-CD19;  
XX anti-CD3; 88.  
XX Mus sp.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX CDS 28..894  
XX FT /\*tag= a  
XX FT /product= "mutant OKT3 with anti-CD19 insert"  
XX CDS 922..1788  
XX FT /\*tag= b  
XX FT /product= "mutant OKT3 with anti-CD3 insert"  
XX DE19721700-Cl.  
XX 19-NOV-1998.  
XX 23-MAY-1997; 97DE-01021700.  
XX 23-MAY-1997; 97DE-01021700.  
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX Kipriyanov S, Little M, Moldenhauer G;  
XX WPI; 1998-596150/51.  
XX P-PSDB; AAW82317, AAW82482.

Monoclonal antibody OKT3 with point-mutation - where cysteine is replaced by another polar amino acid, useful for controlling transplant rejection, and in tumour diagnostics and therapy.  
FS Disclosure; Fig 3; 8pp; German.  
XX This sequence encodes a monoclonal antibody (MAb) diabody derived from OKT3 with a point-mutation where Cys at position H100A is replaced with another polar amino acid, in this example Ser. The diabody encodes two OKT3 proteins, one which has anti-CD19 activity and one with anti-CD3 activity. The MAb is used in lowering or eliminating the transplant rejection in an organ recipient and for diagnostic methods for tumours and tumour therapy  
SQ Sequence 1794 BP; 480 A; 468 C; 450 G; 396 T; 0 U; 0 Other;

Query Match 73.8%; Score 261.2; DB 2; Length 1794;

Best Local Similarity 86.1%; Pred. No. 3.8e-68;  
Matches 303; Conservative 0; Mismatches 43; Indels 6; Gaps 1;  
Qy 9 GCTGCTCGAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCCTCAGTGAAGATGTCTCTG 68  
Db 99 GCAGCTCGAGCAGTCTGGGGCTGAACCTGGCAAGACCTGGGGCCTCAGTGAAGATGTCTCTG 158  
Qy 69 CAAGGGCTTCTGGGCTACACCTTTTACTAACTACTGGATTCACTGGGTGAAACAGAGGCCCTGG 128  
Db 159 CAAGGGCTTCTGGGCTACACCTTTTACTAGGTACACGATGCATGGGTAAACAGAGGCCCTGG 218  
Qy 129 ACAGGGTCTGAATGGATTGGATACATTAATCCTGGCCACTGGTTCACACTTCTTACAATCA 188  
Db 219 ACAGGGTCTGGATGGATTGGATACATTAATCCTAGCCGTGGTTATATACTAATTAACAATCA 278  
Qy 189 GGACTTTTCAGGACAGGCGCACTTTTGACCGCAGACAAAGTCCTCCACACAGCCTACATGCA 248  
Db 279 GAAGTTCAAGGACAGGCGCACTTACGATACAGACAAATCCTCCAGCAGCCTACATGCA 338  
Qy 249 GCTGACCAAGCCTGACATCTGAGGACTCTTCACTTATTACTGTGCAAGA-----GAGGG 302  
Db 339 ACTGAGCAGCCTGACATCTGAGGACTCTGAGTCTTATTACTGTGCAAGATATTATGATGA 398  
Qy 303 GTAGCAGGGTTTGACTCTCGGGCCAGGACCACTCTCAGTCTCCTCA 354  
Db 399 TCATTACAGCCTTGACTACTCGGGCCAGGACCACTCTCAGTCTCCTCA 450

Search completed: June 4, 2005, 07:18:25  
Job time : 709.124 Secs







TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1797 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1797  
OTHER INFORMATION: /product= "Anti-Tac(Fv)-PE40"  
US-08-463-163-2

Query Match 75.5%; Score 267.2; DB 1; Length 1797;  
Best Local Similarity 88.1%; Pred. No. 6.1e-77;  
Matches 303; Conservative 0; Mismatches 38; Indels 3; Gaps 1;  
QY 10 CTGCTCGAGCAGCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
Db 10 CAGCTGCGAGCAGCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
QY 70 AAGGCTTCTGGCTACACCTTTACTAACTACTGGAATTCAGTGGGTGAAACAGAGGCTGGA 129  
Db 70 AAGGCTTCTGGCTACACCTTTACTAACTACTGGAATTCAGTGGGTGAAACAGAGGCTGGA 129  
QY 130 CAGGCTTGAATGATGATGATACATTAATCTCGCCACTGGTTCACCTTCTTACAATCAG 189  
Db 130 CAGGCTTGAATGATGATGATACATTAATCTCGCCACTGGTTCACCTTCTTACAATCAG 189  
QY 190 GACTTTCAGGACAGGCGCACTTGGACCGGACAGTCTCCACACAGCCCTACATGCA 249  
Db 190 AAGTTCAAGGACAGGCGCACTTGGACCGGACAGTCTCCACACAGCCCTACATGCA 249  
QY 250 CTGACGAGCTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGAGAGGGGATGCA 309  
Db 250 CTGACGAGCTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGAGAGGGGATGCA 309  
QY 310 GGGTTTGACTCTCTGGGGCAAGGCAAGCCACTCTCAGAGTCTCTC 353  
Db 307 GTCTTTGACTACTGGGGCAAGGCAAGCCACTCTCAGAGTCTCTC 350

RESULT 4  
PCT-US93-11611-3  
Sequence 3, Application PC/TUS9311611  
GENERAL INFORMATION:  
APPLICANT: Co. Man Sung  
APPLICANT: Landolfi, Nicholas F.  
TITLE OF INVENTION: Humanized Antibodies Reactive with CD18  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/11611  
FILING DATE: 30-NOV-1993  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/983,949  
FILING DATE: 01-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.

REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-45  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 408 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..408  
PCT-US93-11611-3  
Query Match 75.3%; Score 266.6; DB 5; Length 408;  
Best Local Similarity 85.8%; Pred. No. 4.7e-77;  
Matches 296; Conservative 0; Mismatches 49; Indels 0; Gaps 0;  
QY 10 CTGCTCGAGCAGCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
Db 64 CAGCTTTCAGCAGTCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 123  
QY 70 AAGGCTTCTGGCTACACCTTTACTAACTACTGGAATTCAGTGGGTGAAACAGAGGCTGGA 129  
Db 124 AAGGCTTCTGGCTACACCTTTACTAGTTACTTGTGATGCACTGGGTGAAACAGAGGCTGGA 183  
QY 130 CAGGCTTGAATGATGATGATACATTAATCTCGCCACTGGTTCACCTTCTTACAATCAG 189  
Db 184 CAGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 243  
QY 190 GACTTTCAGGACAGGCGCACTTGGACCGGACAGTCTCCACACAGCCCTACATGCA 249  
Db 244 AACTTTCAGGACAGGCGCACTTGGACCGGACAGTCTCCACACAGCCCTACATGCA 303  
QY 250 CTGACGAGCTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGAGAGGGGATGCA 309  
Db 304 CTGACGAGCTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGAGAGGGGATGCA 363  
QY 310 GGGTTTGACTCTCTGGGGCAAGGCAAGCCACTCTCAGAGTCTCTC 354  
Db 364 CTTATGAGTACTTGGGGTCAAGGAACTCAGTCAAGGCTCTCTC 408

RESULT 5  
US-09-065-059-16  
Sequence 16, Application US/09065059  
Patent No. 6068841  
GENERAL INFORMATION:  
APPLICANT: SEINO, Ken-ichi  
APPLICANT: KAYAGAKI, No. 6068841uhiko  
APPLICANT: YAGITA, Hideo  
APPLICANT: OKUMURA, Ko  
APPLICANT: NAKATA, Motomi  
TITLE OF INVENTION: THERAPEUTIC AGENT FOR HEPATITIS  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDermott, Will & Emery  
STREET: 99 Canal Center Plaza  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/065,059  
FILING DATE:

```
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Bucca Ph.D., Daniel
; REGISTRATION NUMBER: P-42,368
; REFERENCE/DOCKET NUMBER: 50356-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-518-5100
; TELEFAX: 703-684-1124
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
US-09-065-059-16

Query Match          74.6%; Score 264.2; DB 3; Length 351;
Best Local Similarity 86.8%; Pred. No. 2.7e-76;
Matches 303; Conservative 0; Mismatches 43; Indels 3; Gaps 1;

Qy 9 GCTGCTCAGCAGCCTGGGGCTGAACTGGCAAAACCTGGGGCCCTCAGTGAAGATGTCTCTG 68
Db 3 GCAGCTCGAGAGTCTGGGGCTGAACTGGCAAAACCTGGGGCCCTCAGTGAAGATGTCTCTG 62

Qy 69 CAAGGCTTCTGGCTACACCTTTTACTAACTACTGGATTCACTGGGTGAAACAGAGGCTCTG 128
Db 63 CAAGGCTTCTGGCTACACCTTTTACTAACTACTGGATTCACTGGGTGAAACAGAGGCTCTG 122

Qy 129 ACAGGGTCTGAATGGATTGGATACATTAATCTTCCACCTGGTCCACTCTTTACAATCA 188
Db 123 ACAGGGTCTGAATGGATTGGATACATTAATCTTCCACCTGGTCCACTCTTTACAATCA 182

Qy 189 GCATTTTCAGGACAGGCGCCACTTTGACCGCAGACAAGTCTCCACACAGCCTACATGCA 248
Db 183 GAAGTTCAGGACAGGCGCCACTTTGACCGCAGACAAGTCTCCACACAGCCTACATGCA 242

Qy 249 GCTACACGAGCTGACATCTGAGGACTCTTCACTCTTATTTACTGTGCAAGAGAGGG---GTA 305
Db 243 ACTAATCAGCCTGACATCTGAGGACTCTTCACTCTTATTTACTGTGCAAGAGAGGGTAATTA 302

Qy 306 CGACGGGTTTCACTCTTGGGGCCCAAGGACACCACTCTCACAGTCTCTCTCA 354
Db 303 CTACTACTTTGACTACTCTGGGGCCCAAGGACACCACTCTCACAGTCTCTCTCA 351
```

## RESULT 6

```
US-08-273-146-58
; Sequence 58, Application US/08273146
; Patent No. 5955885
```

## GENERAL INFORMATION:

```
; APPLICANT: Smith, Rodger
; APPLICANT: McCafferty, John
; APPLICANT: Chiswell, David
; APPLICANT: Darsley, Michael J.
; APPLICANT: Fitzgerald, Kevin
; APPLICANT: Kenten, John H.
; APPLICANT: Martin, Mark T.
; APPLICANT: Titmas, Richard C.
; APPLICANT: Williams, Richard O.
; TITLE OF INVENTION: The isolation and production of
; TITLE OF INVENTION: Catalytic Antibodies using Phage Technology
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IGEN, Inc.
; STREET: 1530 East Jefferson St.
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20852
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/273,146
```

```
; FILING DATE: 14-JUL-1994
```

```
; CLASSIFICATION: 435
```

```
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Ryan, John W.
```

```
; REGISTRATION NUMBER: 33,771
```

```
; REFERENCE/DOCKET NUMBER: 09000
```

```
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 301-984-8000
```

```
; TELEFAX: 301-230-0158
```

```
; INFORMATION FOR SEQ ID NO: 58:
```

```
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 354 base pairs
```

```
; TYPE: nucleic acid
```

```
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
```

```
; FEATURE:
```

```
; NAME/KEY: CDS
```

```
; LOCATION: 1..339
```

```
US-08-273-146-58
```

```
Query Match          74.6%; Score 264.2; DB 2; Length 354;
```

```
Best Local Similarity 86.8%; Pred. No. 2.7e-76;
```

```
Matches 303; Conservative 0; Mismatches 43; Indels 3; Gaps 1;
```

```
Qy 9 GCTGCTCAGCAGCCTGGGGCTGAACTGGCAAAACCTGGGGCCCTCAGTGAAGATGTCTCTG 68
Db 6 GCAGCTCGAGAGTCTGGGGCTGAACTGGCAAAACCTGGGGCCCTCAGTGAAGATGTCTCTG 65

Qy 69 CAAGGCTTCTGGCTACACCTTTTACTAACTACTGGATTCACTGGGTGAAACAGAGGCTCTG 128
Db 66 CAAGGCTTCTGGCTACACCTTTTACTAACTACTGGATTCACTGGGTGAAACAGAGGCTCTG 125

Qy 129 ACAGGGTCTGAATGGATTGGATACATTAATCTTCCACCTGGTCCACTCTTTACAATCA 188
Db 126 ACAGGGTCTGAATGGATTGGATACATTAATCTTCCACCTGGTCCACTCTTTACAATCA 185

Qy 189 GCATTTTCAGGACAGGCGCCACTTTGACCGCAGACAAGTCTCCACACAGCCTACATGCA 248
Db 186 GAAGTTCAGGACAGGCGCCACTTTGACCGCAGACAAGTCTCCACACAGCCTACATGCA 245

Qy 249 GCTACACGAGCTGACATCTGAGGACTCTTCACTCTTATTTACTGTGCAAGAGAGGG---GGGTA 305
Db 246 ACTGAGCAGCTGACATCTGAGGACTCTTCACTCTTATTTACTGTGCAAGAGAGGGTAATTA 305

Qy 306 CGACGGGTTTCACTCTTGGGGCCCAAGGACACCACTCTCACAGTCTCTCTCA 354
Db 306 CTATGCTATGAGTACTCTGGGGCCCAAGGACACCACTCTCACAGTCTCTCTCA 354
```

## RESULT 7

```
US-08-116-247-6
```

```
; Sequence 6, Application US/08116247
```

```
; Patent No. 5929212
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Jolliffe, Linda K.
```

```
; APPLICANT: Zivin, Robert A.
```

```
; APPLICANT: Adair, John R.
```

```
; APPLICANT: Athwal, Diljeet S.
```

```
; TITLE OF INVENTION: CD3 Specific Recombinant Antibody
```

```
; NUMBER OF SEQUENCES: 29
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Woodcock Wasburn Kurtz Mackiewicz & No. 5929212rlis
```

```
; STREET: One Liberty Place - 46th Floor
```

```
; CITY: Philadelphia
```

```
; STATE: PA
```

```
; COUNTRY: USA
```

```
; ZIP: 19103
```

```
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/116.247  
FILING DATE: 07/08/116.247  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/743,377  
FILING DATE: 10-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Paintin, Francis A.  
REGISTRATION NUMBER: 19,386  
REFERENCE/DOCKET NUMBER: CARP-0011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1570 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 41..1444  
US-08-116-247-6

Query Match 73.5%; Score 260.2; DB 2; Length 1570;  
Best Local Similarity 86.0%; Pred. No. 1.1e-74;  
Matches 302; Conservative 0; Mismatches 43; Indels 6; Gaps 1;  
QY 10 CTGCTCGAGCAGCCTGGGGCTGAACCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
DB 104 CAGCTGCAGCAGTCTGGGGCTGAACTGGCAAGACCTGGGGCTCAGTGAAGATGTCCTGC 163  
QY 70 AAGGCTTCTGGCTACACCTTTACTTAATCTGGAATCTGCTGGTGAACAGAGGCTTGA 129  
DB 164 AAGGCTTCTGGCTACACCTTTACTAGGTACAGATGCACTGGTGAACAGAGGCTTGA 223  
QY 130 CAGGCTCTGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 189  
DB 224 CAGGCTCTGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 283  
QY 190 GACTTTTCAGGACAGGCGCACTTTGACCGCAGACAGCAAGTCTCCACACAGCCTTACATGCA 249  
DB 284 AAGTTCAAGGACAGGCGCACTTTGACCGCAGACAGCAAGTCTCCACACAGCCTTACATGCA 343  
QY 250 CTGACCGCCTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGA-----GAGGG 303  
DB 344 CTGACCGCCTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGATATTATGATGAT 403  
QY 304 TACGACGGGTTGACTCTCTGGGGCCAAAGGACCACTCTCAGAGTCTCCTCA 354  
DB 404 CATTACTGCTTGAATCTGAGGACTCTGAGGACTCTGAGTCTATTACTGTGCAAGATATTATGATGAT 454

RESULT 8  
US-09-348-224-6  
Sequence 6, Application US/09348224  
Patent No. 6750325  
GENERAL INFORMATION:  
APPLICANT: Joliffe, Linda Kay  
APPLICANT: Zivin, Robert Allan  
APPLICANT: Adair, John Robert  
APPLICANT: Athwal, Diljeet Singh  
TITLE OF INVENTION: CD3 Specific Recombinant Antibody  
FILE REFERENCE: CARP0066  
CURRENT APPLICATION NUMBER: US/09/348,224  
CURRENT FILING DATE: 1999-07-06  
EARLIER APPLICATION NUMBER: 08/116,247  
EARLIER FILING DATE: 1993-09-03  
EARLIER APPLICATION NUMBER: 07/743,377

EARLIER FILING DATE: 1991-10-04  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 6  
LENGTH: 1570  
TYPE: DNA  
ORGANISM: Mouse  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (41)..(1444)  
US-09-348-224-6  
Query Match 73.5%; Score 260.2; DB 4; Length 1570;  
Best Local Similarity 86.0%; Pred. No. 1.1e-74;  
Matches 302; Conservative 0; Mismatches 43; Indels 6; Gaps 1;  
QY 10 CTGCTCGAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
DB 104 CAGCTGCAGCAGTCTGGGGCTGAACTGGCAAGACCTGGGGCTCAGTGAAGATGTCCTGC 163  
QY 70 AAGGCTTCTGGCTACACCTTTACTTAATCTGGAATCTGCTGGTGAACAGAGGCTTGA 129  
DB 164 AAGGCTTCTGGCTACACCTTTACTAGGTACAGATGCACTGGTGAACAGAGGCTTGA 223  
QY 130 CAGGCTCTGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 189  
DB 224 CAGGCTCTGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 283  
QY 190 GACTTTTCAGGACAGGCGCACTTTGACCGCAGACAGCAAGTCTCCACACAGCCTTACATGCA 249  
DB 284 AAGTTCAAGGACAGGCGCACTTTGACCGCAGACAGCAAGTCTCCACACAGCCTTACATGCA 343  
QY 250 CTGACCGCCTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGA-----GAGGG 303  
DB 344 CTGACCGCCTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGATATTATGATGAT 403  
QY 304 TACGACGGGTTGACTCTCTGGGGCCAAAGGACCACTCTCAGAGTCTCCTCA 354  
DB 404 CATTACTGCTTGAATCTGAGGACTCTGAGGACTCTGAGTCTATTACTGTGCAAGATATTATGATGAT 454

RESULT 9  
US-08-303-569B-6  
Sequence 6, Application US/08303569B  
Patent No. 5859205  
GENERAL INFORMATION:  
APPLICANT: Adair, John R.  
APPLICANT: Athwal, Diljeet S.  
APPLICANT: Emtage, John S.  
TITLE OF INVENTION: Humanised Antibodies  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5859205ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/303,569B  
FILING DATE: 07-SEP-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Trujillo, Doreen Yanko  
REGISTRATION NUMBER: 35,719  
REFERENCE/DOCKET NUMBER: CARP-0032  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100

```
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1570 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 41..1444
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 98..1444
; US-08-303-569B-6

Query Match      73.1%; Score 258.6; DB 2; Length 1570;
Best Local Similarity 85.8%; Pred. No. 3.7e-74;
Matches 301; Conservative 0; Mismatches 44; Indels 6; Gaps 1;

Qy 10 CTGCTCGAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69
Db 104 CAGCTGCAGCAGTCTGGGGCTGAACCTGGCAAGACTGGGGCTCAGTGAAGATGTCCTGC 163
Qy 70 AAGGCTTCTGGCTACACCTTTTACTAACTACTGGAATTCACCTGGGTGAACAGAGCCCTGGA 129
Db 164 AAGGCTTCTGGCTACACCTTTTACTAGGTACACGATGCACTGGGTAAACAGAGCCCTGGA 223
Qy 130 CAGGCTTGAATGGATGGATACATTAATCTCCCACTGGTCCACTCTTCAATCAG 189
Db 224 CAGGCTTGAATGGATGGATACATTAATCTAGCCGTGGTTACTTAATTAATCAG 283
Qy 190 GACTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCCACACAGCCCTACATGCA 249
Db 284 AAGTTCAAGGACAGGGCCACTTTGACCGCAGACAATCTCCACAGCCCTACATGCA 343
Qy 250 CTGACAGCCTGACATCTGAGGACTCTTTCAGTCTTACTTGTGCAAGA-----GAGGGG 303
Db 344 CTGACAGCCTGACATCTGAGGACTCTGCACTCTTACTTGTGCAAGATATTATGATGAT 403
Qy 304 TAGCAGGGTTGACTCTCTGGGGCCAGGACCACTCTCAGAGTCTCTCA 354
Db 404 CATTACTGCTTGAAGTCTGAGGCTGAGGACCACTCTCAGAGTCTCTCA 454

RESULT 10
US-09-795-515-6
; Sequence 6, Application US/09795515
; Patent No. 663297
; GENERAL INFORMATION:
; APPLICANT: Adair, John R.
; APPLICANT: Athwal, Diljeet S.
; APPLICANT: Emtage, John S.
; TITLE OF INVENTION: Humanised Antibodies
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6632927ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/795,515
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/846,658
```

```
; FILING DATE: 01-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yanko
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CARP-0057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1570 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 41..1444
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 98..1444
; US-09-795-515-6

Query Match      73.1%; Score 258.6; DB 4; Length 1570;
Best Local Similarity 85.8%; Pred. No. 3.7e-74;
Matches 301; Conservative 0; Mismatches 44; Indels 6; Gaps 1;

Qy 10 CTGCTCGAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69
Db 104 CAGCTGCAGCAGTCTGGGGCTGAACCTGGCAAGACTGGGGCTCAGTGAAGATGTCCTGC 163
Qy 70 AAGGCTTCTGGCTACACCTTTTACTAACTACTGGAATTCACCTGGGTGAACAGAGCCCTGGA 129
Db 164 AAGGCTTCTGGCTACACCTTTTACTAGGTACACGATGCACTGGGTAAACAGAGCCCTGGA 223
Qy 130 CAGGCTTGAATGGATGGATACATTAATCTCCCACTGGTCCACTCTTCAATCAG 189
Db 224 CAGGCTTGAATGGATGGATACATTAATCTAGCCGTGGTTACTTAATTAATCAG 283
Qy 190 GACTTTCAGGACAGGGCCACTTTGACCGCAGACAAGTCTCCACACAGCCCTACATGCA 249
Db 284 AAGTTCAAGGACAGGGCCACTTTGACCGCAGACAATCTCCACAGCCCTACATGCA 343
Qy 250 CTGACAGCCTGACATCTGAGGACTCTTTCAGTCTTACTTGTGCAAGA-----GAGGGG 303
Db 344 CTGACAGCCTGACATCTGAGGACTCTGCACTCTTACTTGTGCAAGATATTATGATGAT 403
Qy 304 TAGCAGGGTTGACTCTCTGGGGCCAGGACCACTCTCAGAGTCTCTCA 354
Db 404 CATTACTGCTTGAAGTCTGAGGCTGAGGACCACTCTCAGAGTCTCTCA 454

RESULT 11
US-08-800-198-1
; Sequence 1, Application US/08800198
; Patent No. 5942602
; GENERAL INFORMATION:
; APPLICANT: WELLS, WINFRIED S.
; APPLICANT: SCHMIDT, MATHIAS
; APPLICANT: VAKALOPOULOU, EVANGELIA
; APPLICANT: SCHNEIDER, DOUGLAS
; TITLE OF INVENTION: GROWTH FACTOR RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD. SUITE 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: US
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/800,198  
FILING DATE: 13-FEB-1997  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: SCH 1576  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 357 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-800-198-1

Query Match 72.6%; Score 257; DB 2; Length 357;  
Best Local Similarity 85.5%; Pred. No. 6.1e-74;  
Matches 300; Conservative 0; Mismatches 45; Indels 6; Gaps 1;

QY 9 GCTGCTGAGCAGCCTGGGGCTGAACTGCGCAAAACCTGGGGCTCAGTGAAGATGTCCTG 68  
DB 6 GCAGCTGCAGGAGTCTGGGGCTGAACTGCGCAAAACCTGGGGCTCAGTGAAGATGTCCTG 65

QY 69 CAAGGCTTCTGGCTACACCTTTTACTTAACCTACTGGAATTCACCTGGGTGAAACAGAGGCCTGG 128  
DB 66 CAAGGCTTCTGGCTACACCTTTTACTTAACCTACTGGAATTCACCTGGGTGAAACAGAGGCCTGG 125

QY 129 ACAGGCTCTGAAATGGATTGGATACATTAATTCCTGCCACTGTTCCACTTCTTACAATCA 188  
DB 126 ACAGGCTCTGAAATGGATTGGATACATTAATTCCTGCCACTGTTCCACTTCTTACAATCA 185

QY 189 GGACTTTTCAGACAGAGGCCTTTTACCGCGAGACAGTCTCTCCACAGCAGCTACATGCA 248  
DB 186 GAAGTTCAAGGACCAAGGCCCATTTGACTGCGACAGCAAAATCCTCCAGCAGCCTACATGCA 245

QY 249 GCTGACAGCCTGACATCTGAGGACTCTTCAAGTCTTACTTACTTGTGCAAGAGGGGTACGA 308  
DB 246 ACTGAGCGGCTGACATCTGAGGACTCTGAGTCTTACTTACTTGTGCAAGAGGGGTACTA 305

QY 309 CGG-----GTTTGAATCTCTGGGGCCAAAGGACCACTCTCACAGTCTCTC 353  
DB 306 CGGCTAGACTTTGTTTACTGGGGCCAAAGGACCACTCTCACAGTCTCTC 356

RESULT 12  
US-09-296-595-1  
Sequence 1, Application US/09296595A  
Patent No. 6129915  
GENERAL INFORMATION:  
APPLICANT: WELLS, WINFRIED S.  
APPLICANT: SCHMIDT, MATHIAS  
APPLICANT: VAKALOPULOU, EVANGELIA  
APPLICANT: SCHNEIDER, DOUGLAS  
TITLE OF INVENTION: GROWTH FACTOR RECEPTOR ANTIBODIES  
FILE REFERENCE: SCH-1576 D1  
CURRENT APPLICATION NUMBER: US/09/296,595A  
CURRENT FILING DATE: 1999-04-23  
EARLIER APPLICATION NUMBER: 08/800,198  
EARLIER FILING DATE: 1997-02-13  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 357  
TYPE: DNA

ORGANISM: Murine sp.  
US-09-296-595-1

Query Match 72.6%; Score 257; DB 3; Length 357;  
Best Local Similarity 85.5%; Pred. No. 6.1e-74;  
Matches 300; Conservative 0; Mismatches 45; Indels 6; Gaps 1;

QY 9 GCTGCTGAGCAGCCTGGGGCTGAACTGCGCAAAACCTGGGGCTCAGTGAAGATGTCCTG 68  
DB 6 GCAGCTGCAGGAGTCTGGGGCTGAACTGCGCAAAACCTGGGGCTCAGTGAAGATGTCCTG 65

QY 69 CAAGGCTTCTGGCTACACCTTTTACTTAACCTACTGGAATTCACCTGGGTGAAACAGAGGCCTGG 128  
DB 66 CAAGGCTTCTGGCTACACCTTTTACTTAACCTACTGGAATTCACCTGGGTGAAACAGAGGCCTGG 125

QY 129 ACAGGCTCTGAAATGGATTGGATACATTAATTCCTGCCACTGTTCCACTTCTTACAATCA 188  
DB 126 ACAGGCTCTGAAATGGATTGGATACATTAATTCCTGCCACTGTTCCACTTCTTACAATCA 185

QY 189 GGACTTTTCAGACAGAGGCCTTTTACCGCGAGACAGTCTCTCCACAGCAGCTACATGCA 248  
DB 186 GAAGTTCAAGGACCAAGGCCCATTTGACTGCGACAGCAAAATCCTCCAGCAGCCTACATGCA 245

QY 249 GCTGACAGCCTGACATCTGAGGACTCTTCAAGTCTTACTTACTTGTGCAAGAGGGGTACGA 308  
DB 246 ACTGAGCGGCTGACATCTGAGGACTCTGAGTCTTACTTACTTGTGCAAGAGGGGTACTA 305

QY 309 CGG-----GTTTGAATCTCTGGGGCCAAAGGACCACTCTCACAGTCTCTC 353  
DB 306 CGGCTAGACTTTGTTTACTGGGGCCAAAGGACCACTCTCACAGTCTCTC 356

RESULT 13  
US-08-482-882-77  
Sequence 77, Application US/08482882  
Patent No. 5773218  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Vazeux, Rosemay  
TITLE OF INVENTION: ICAM-Related Materials and Methods  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,882  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,754  
FILING DATE:  
APPLICATION NUMBER: US 08/102,852  
FILING DATE: 05-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,266  
FILING DATE: 22-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/894,061  
FILING DATE: 05-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/889,724  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,689

```
; FILING DATE: 27-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5773218and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32178
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-482-882-77

Query Match 72.6%; Score 257; DB 1; Length 422;
Best Local Similarity 85.5%; Pred. No. 6.6e-74;
Matches 300; Conservative 0; Mismatches 45; Indels 6; Gaps 1;

Qy 10 CTGCTCGAGCAGCTGGGGCTGAACTGGCAAACTGGGGCTCAGTGAAGATGTCCTGC 69
Db 72 CAGCTTCAGCAGCTGGGGCTGAACTGGCAAACTGGGGCTCAGTGAAGATGTCCTGC 131
Qy 70 AAGGCTTCGCTACACCTTTACTAACTACTCGGATTCACCTGGGTGAACAGAGCCCTGGA 129
Db 132 AAGGCTTCGCTACACCTTTACTGTTTACTGGATGCACTGGGTAAACAGAGCCCTGGA 191
Qy 130 CAGGCTCGAAATGGATTGGATACATTAATCCTGCCACTGGTTCACCTTTACAATCAG 189
Db 192 CAGGCTTCAGAAATGGATTGGATACATTAATCCTAACACTGATTTACTTGAGTACAATCAG 251
Qy 190 GACTTTCAGGACAGGCGCACTTTCACCGCAGACAAGTCCCTCCACAGCCTCATGCGAG 249
Db 252 AGGTTCCAGGACAAGGCGCACATTCAGTCGACAAAATCCTCCAGCAGAGCCTCATGCAA 311
Qy 250 CTGACCAAGCCTGACATCTGAGGACTCTTCACTTACTTACTTGTGCAAGAGAGGGG- 303
Db 312 CTGAGCAGCCTGACATCTGAGGACTCTGAGTCTTATTACTTGTGCAAGATGGGGGGTAA 371
Qy 304 TAGCAGGGTTTGACTCTCTGGGGCAAGGCAACCACTCTCACAGTCTCTCA 354
Db 372 TCCTATGTTTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCA 422

RESULT 14
US-08-483-389-77
; Sequence 77, Application US/08483389
; Patent No. 5811517
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Vazeux, Rosemay
; TITLE OF INVENTION: ICAM-RELATED PROTEIN
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483.389
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Suh, Young J.
; REGISTRATION NUMBER: P-41,337
; REFERENCE/DOCKET NUMBER: 27866/32760
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: (312) 474-6600
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-483-389-77

Query Match 72.6%; Score 257; DB 1; Length 422;
Best Local Similarity 85.5%; Pred. No. 6.6e-74;
Matches 300; Conservative 0; Mismatches 45; Indels 6; Gaps 1;

Qy 10 CTGCTCGAGCAGCCTGGGGCTGAACTGGCAAACTGGGGCTCAGTGAAGATGTCCTGC 69
Db 72 CAGCTTCAGCAGCTGGGGCTGAACTGGCAAACTGGGGCTCAGTGAAGATGTCCTGC 131
Qy 70 AAGGCTTCGCTACACCTTTACTAACTACTCGGATTCACCTGGGTGAACAGAGCCCTGGA 129
Db 132 AAGGCTTCGCTACACCTTTACTGTTTACTGGATGCACTGGGTAAACAGAGCCCTGGA 191
Qy 130 CAGGCTCGAAATGGATTGGATACATTAATCCTGCCACTGGTTCACCTTTACAATCAG 189
Db 192 CAGGCTTCAGAAATGGATTGGATACATTAATCCTAACACTGATTTACTTGAGTACAATCAG 251
Qy 190 GACTTTCAGGACAGGCGCACTTTCACCGCAGACAAGTCCCTCCACAGCCTCATGCGAG 249
Db 252 AGGTTCCAGGACAAGGCGCACATTCAGTCGACAAAATCCTCCAGCAGAGCCTCATGCAA 311
Qy 250 CTGACCAAGCCTGACATCTGAGGACTCTTCACTTACTTACTTGTGCAAGAGAGGGG- 303
Db 312 CTGAGCAGCCTGACATCTGAGGACTCTGAGTCTTATTACTTGTGCAAGATGGGGGGTAA 371
Qy 304 TAGCAGGGTTTGACTCTCTGGGGCAAGGCAACCACTCTCACAGTCTCTCA 354
Db 372 TCCTATGTTTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCA 422

RESULT 15
US-08-487-113D-77
; Sequence 77, Application US/08487113D
; Patent No. 5837822
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Vazeux, Rosemay
; TITLE OF INVENTION: ICAM-Related Materials and Methods
; NUMBER OF SEQUENCES: 120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
```

STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,113D  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,754  
FILING DATE: 05-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/102,852  
FILING DATE: 05-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/009,266  
FILING DATE: 22-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/894,061  
FILING DATE: 05-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/889,724  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,689  
FILING DATE: 27-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5837822and, Greta E.  
REGISTRATION NUMBER: 35,502  
REFERENCE/DOCKET NUMBER: 32744  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 77:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 422 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-487-113D-77

Query Match 72.6%; Score 257; DB 2; Length 422;  
Best Local Similarity 85.5%; Pred. No. 6,6e-74;  
Matches 300; Conservative 0; Mismatches 45; Indels 6; Gaps 1;  
Qy 10 CTGCTCGAGCAGCCTGGGGCTGAACCTGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
Db 72 CAGCTTCAGCAGTCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 131  
Qy 70 AAGGCTTCGTGCTACACCTTTACTAACTACTGGATTCACTGGGTGAACAGAGGCTTGA 129  
Db 132 AAGGCTTCGTGCTACACCTTTACTGTTTACTGGATGCACTGGGTAAACAGAGGCTTGA 191  
Qy 130 CAGGCTCTGAATGATGGATACATTAATCTGCACCTGGTTCACCTCTTACAAATCAG 189  
Db 192 CAGGCTCTGAATGATGGATACATTAATCTGCACCTGGTTCACCTCTTACAAATCAG 251  
Qy 190 GACTTTCAGGACAGGGCCACTTTGACCGCAGACAGTCTCTCCACACAGCCTACATGCAG 249  
Db 252 AGGTTCCAGGACAGGGCCACTTTGACCGCAGACAGTCTCTCCACACAGCCTACATGCA 311  
Qy 250 CTGACGACCTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGAGAGGGG-----303  
Db 312 CTGACGACCTGACATCTGAGGACTCTGCACTCTATTACTGTGCAAGATGGGGGGTAAC 371  
Qy 304 TACGACGGGTTTGACTCTCTGGGGCAAGGCACTCTCAGATCTCCTCA 354  
US-08-487-113D-77

Db 372 TCCTATGGTTTGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCA 422

Search completed: June 4, 2005, 10:27:48  
Job time : 214.556 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 08:29:26 ; Search time 854.914 Seconds  
(without alignments)  
2545.505 Million cell updates/sec

Title: US-10-089-452-1  
Perfect score: 354  
Sequence: 1 gaggtcgagctgctcgagca.....ccactctcacagtctctcca 354

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	268.2	75.8	406	15	US-10-150-762-85 Sequence 85, Appl
2	268.2	75.8	406	15	US-10-150-762-87 Sequence 87, Appl
3	268.2	75.8	406	15	US-10-244-821-85 Sequence 85, Appl
4	266.8	75.4	348	14	US-10-108-511-6 Sequence 6, Appli
5	266.8	75.4	348	19	US-10-482-532-6 Sequence 6, Appli
6	261.2	73.8	1817	19	US-10-489-626-4 Sequence 4, Appli
7	261.2	73.8	1817	19	US-10-489-626-11 Sequence 11, Appl
8	260.2	73.5	357	18	US-10-682-845-43 Sequence 43, Appl
9	260.2	73.5	417	18	US-10-682-845-41 Sequence 41, Appl
10	260.2	73.5	1570	18	US-10-682-845-39 Sequence 39, Appl
11	260.2	73.5	10511	14	US-10-059-261-109 Sequence 109, App

12	260.2	73.5	10511	18	US-10-627-649-109 Sequence 109, App
13	259.6	73.3	1817	19	US-10-489-626-6 Sequence 6, Appli
14	258.6	73.1	1570	10	US-09-795-515-6 Sequence 6, Appli
15	258.6	73.1	1570	17	US-10-704-352-6 Sequence 6, Appli
16	258.6	73.1	1570	18	US-10-703-963-6 Sequence 6, Appli
17	257.2	72.7	1479	18	US-10-682-845-82 Sequence 82, Appl
18	257	72.6	422	9	US-09-753-436-77 Sequence 77, Appl
19	257	72.6	422	16	US-10-163-942-77 Sequence 77, Appl
20	257	72.6	422	18	US-10-745-115-77 Sequence 77, Appl
21	255.6	72.2	1241	17	US-10-362-591-1 Sequence 60, Appl
22	255.6	72.2	1479	18	US-10-682-845-60 Sequence 62, Appl
23	255.6	72.2	1479	18	US-10-682-845-62 Sequence 64, Appl
24	255.6	72.2	1479	18	US-10-682-845-64 Sequence 66, Appl
25	255.6	72.2	1479	18	US-10-682-845-66 Sequence 68, Appl
26	255.6	72.2	1479	18	US-10-682-845-68 Sequence 80, Appl
27	255.6	72.2	1479	18	US-10-682-845-80 Sequence 86, Appl
28	255.6	72.2	1479	18	US-10-682-845-88 Sequence 88, Appl
29	255.6	72.2	1560	19	US-10-805-177-110 Sequence 110, Appl
30	255.6	72.2	1574	16	US-10-168-809-21 Sequence 21, Appl
31	255.6	72.2	1635	19	US-10-805-177-112 Sequence 112, App
32	255.4	72.1	360	19	US-10-879-994-97 Sequence 97, Appl
33	255.4	72.1	729	19	US-10-879-994-9 Sequence 9, Appli
34	255.4	72.1	729	19	US-10-610-452-9 Sequence 25, Appl
35	254	71.8	363	10	US-09-948-004-25 Sequence 70, Appl
36	254	71.8	1479	18	US-10-682-845-70 Sequence 72, Appl
37	254	71.8	1479	18	US-10-682-845-72 Sequence 74, Appl
38	254	71.8	1479	18	US-10-682-845-74 Sequence 76, Appl
39	254	71.8	1479	18	US-10-682-845-76 Sequence 78, Appl
40	254	71.8	1479	18	US-10-682-845-84 Sequence 84, Appl
41	254	71.8	1479	18	US-10-682-845-84 Sequence 17, Appl
42	254	71.8	1545	10	US-09-948-004-17 Sequence 3, Appli
43	252.8	71.4	348	9	US-09-741-843-3 Sequence 3, Appli
44	252.8	71.4	348	10	US-09-894-839-3 Sequence 3, Appli
45	252.8	71.4	348	10	US-09-894-839-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-10-150-762-85  
; Sequence 85, Application US/10150762  
; Publication No. US20030103948A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen C.  
; APPLICANT: Graves, Scott S.  
; APPLICANT: Schultz, Joanne E.  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James A.  
; APPLICANT: Reno, Jonh M.  
; APPLICANT: Dearstynne, Erica A.  
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
; FILE REFERENCE: 690022.547C2  
; CURRENT APPLICATION NUMBER: US/10/150,762  
; CURRENT FILING DATE: 2002-05-17  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 85  
; LENGTH: 406  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-150-762-85

Query Match	75.8%	Score 268.2;	DB 15;	Length 406;
Best Local Similarity	88.1%	Pred. No. 1.3e-78;		
Mismatches	304;	Conservative	0;	Mismatches 38; Indels 3; Gaps 1;
Qy	10	CTGCTCGAGCAGCTGGGGCTGAACCTGGGAACTGGGGCTCAGTGAAGATGTCCTGC	69	
Db	64	CAGCTTCAGCAGCTGGGGCTGAACCTGGGAACTGGGGCTCAGTGAAGATGTCCTGC	123	
Qy	70	ANGGCTTCTGGCTACACCTTTACTACTGATTCACTGGTGAACAGAGGCTGGA	129	

Db 124 AAGGCTTCTGGCTACACCTTTTACTAGCTACAGGATGCACTGGTAAACAGAGGCGCTGGA 183  
Qy 130 CAGGCTCTGAATGGATTGGATACATTAACTCTGCCACTGTCTCCACTTCTTACAATCAG 189  
Db 184 CAGGCTCTGAATGGATTGGATATATTAATCTTAGCACTGGGTATCTGAATACAATCAG 243  
Qy 190 GACTTTTCAGGACAGGCGCACTTTGACCGCAGACAAGTCTCTCCACACAGCCTACATGCG 249  
Db 244 AAGTTCAAGGACAAGGCCACATTGACTGCAGACAAATCTCCAGCACAGCCTACATGCAA 303  
Qy 250 CTGACCGCCTGACATCTGAGGACTCTTCACTCTTACTCTGCAAGAGGGGTACGAC 309  
Db 304 CTGACGACGCTGACATTTGAGGACTCTGCACTCTATTACTGTGCAAGAGGGGGG---GGG 360  
Qy 310 GGGTTTGACTCTCTGGGGCAAGGACCACTCTCACAGTCTCTCA 354  
Db 361 GTCTTTGACTACTGGGGCCAAAGAACCACTCTCACAGTCTCTCA 405

## RESULT 2

US-10-150-762-87

; Sequence 87, Application US/10150762

; Publication No. US20030103948A1

; GENERAL INFORMATION:

; APPLICANT: Goshorn, Stephen C.

; APPLICANT: Graves, Scott S.

; APPLICANT: Schultz, Joanne E.

; APPLICANT: Lin, Yukang

; APPLICANT: Sanderson, James A.

; APPLICANT: Reno, John M.

; APPLICANT: Dearstine, Erica A.

; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND

; FILE REFERENCE: 690022.547C2

; CURRENT APPLICATION NUMBER: US/10/150,762

; CURRENT FILING DATE: 2002-05-17

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 87

; LENGTH: 406

; TYPE: DNA

; ORGANISM: Mus musculus

; US-10-150-762-87

Query Match 75.8%; Score 268.2; DB 15; Length 406;  
Best Local Similarity 88.1%; Pred. No. 1.3e-78;  
Matches 304; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

Qy 10 CTGCTCAGCAGCCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
Db 64 CAGCTTCAGCAGTCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 123  
Qy 70 AAGGCTTCTGGCTACACCTTTTACTAACTACTGGATTCTCTGGGTGAACAGAGGCGCTGGA 129  
Db 124 AAGGCTTCTGGCTACACCTTTTACTAGCTACAGGATGCACTGGGTAAACAGAGGCGCTGGA 183  
Qy 130 CAGGCTCTGAATGGATTGGATACATTAACTCTGCCACTGTCTCCACTTCTTACAATCAG 189  
Db 184 CAGGCTCTGGATGGATTGGATATATTAATCTTAGCACTGGGTATCTGAATACAATCAG 243  
Qy 190 GACTTTTCAGGACAGGCGCACTTTGACCGCAGACAAGTCTCTCCACACAGCCTACATGCG 249  
Db 244 AAGTTCAAGGACAAGGCCACATTGACTGCAGACAAATCTCCAGCACAGCCTACATGCAA 303  
Qy 250 CTGACCGCCTGACATCTGAGGACTCTTCACTCTTACTCTGCAAGAGGGGTACGAC 309  
Db 304 CTGACGACGCTGACATTTGAGGACTCTGCACTCTATTACTGTGCAAGAGGGGGG---GGG 360  
Qy 310 GGGTTTGACTCTCTGGGGCAAGGACCACTCTCACAGTCTCTCA 354  
Db 361 GTCTTTGACTACTGGGGCCAAAGAACCACTCTCACAGTCTCTCA 405

## RESULT 3

US-10-244-821-85

; Sequence 85, Application US/10244821

; Publication No. US2003014323A1

; GENERAL INFORMATION:

; APPLICANT: Goshorn, Stephen Charles

; APPLICANT: Graves, Scott Stoll

; APPLICANT: Schultz, Joanne Elaine

; APPLICANT: Lin, Yukang

; APPLICANT: Sanderson, James Allen

; APPLICANT: Reno, John M.

; APPLICANT: Dearstine, Erica A.

; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND

; FILE REFERENCE: 690022.547C3

; CURRENT APPLICATION NUMBER: US/10/244,821

; CURRENT FILING DATE: 2002-09-16

; NUMBER OF SEQ ID NOS: 92

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 85

; LENGTH: 406

; TYPE: DNA

; ORGANISM: Mus musculus

; US-10-244-821-85

Query Match 75.8%; Score 268.2; DB 15; Length 406;  
Best Local Similarity 88.1%; Pred. No. 1.3e-78;  
Matches 304; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

Qy 10 CTGCTCAGCAGCCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
Db 64 CAGCTTCAGCAGTCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 123  
Qy 70 AAGGCTTCTGGCTACACCTTTTACTAACTACTGGATTCTCTGGGTGAACAGAGGCGCTGGA 129  
Db 124 AAGGCTTCTGGCTACACCTTTTACTAGCTACAGGATGCACTGGGTAAACAGAGGCGCTGGA 183  
Qy 130 CAGGCTCTGAATGGATTGGATACATTAACTCTGCCACTGTCTCCACTTCTTACAATCAG 189  
Db 184 CAGGCTCTGGATGGATTGGATATATTAATCTTAGCACTGGGTATCTGAATACAATCAG 243  
Qy 190 GACTTTTCAGGACAGGCGCACTTTGACCGCAGACAAGTCTCTCCACACAGCCTACATGCG 249  
Db 244 AAGTTCAAGGACAAGGCCACATTGACTGCAGACAAATCTCCAGCACAGCCTACATGCAA 303  
Qy 250 CTGACCGCCTGACATCTGAGGACTCTTCACTCTTACTCTGCAAGAGGGGTACGAC 309  
Db 304 CTGACGACGCTGACATTTGAGGACTCTGCACTCTATTACTGTGCAAGAGGGGGG---GGG 360  
Qy 310 GGGTTTGACTCTCTGGGGCCCAAGGACCACTCTCACAGTCTCTCA 354  
Db 361 GTCTTTGACTACTGGGGCCCAAGGACCACTCTCACAGTCTCTCA 405

## RESULT 4

US-10-108-511-6

; Sequence 6, Application US/10108511

; Publication No. US2003001713A1

; GENERAL INFORMATION:

; APPLICANT: Reiter, Yoram

; APPLICANT: Avital, Lev

; TITLE OF INVENTION: METHODS AND PHARMACEUTICAL COMPOSITIONS FOR IMMUNE DECEPTION,

; FILE REFERENCE: 01/23109

; CURRENT APPLICATION NUMBER: US/10/108,511

; CURRENT FILING DATE: 2002-03-29

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 348

; TYPE: DNA

; ORGANISM: Artificial sequence

FEATURE:  
; OTHER INFORMATION: atacvH sequence - a part of the B2M-aTAc (dsFv) construct sequence  
US-10-108-511-6

Query Match 75.4%; Score 266.8; DB 14; Length 348;  
Best Local Similarity 88.3%; Pred. No. 3.6e-78;  
Matches 302; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

QY 13 CTCGAGCAGCTGGGGCTGAACCTGCAAAACCTGGGGCTCAGTGAAGATGTCCTGCAAG 72  
DB 10 CTGACGAGTCTGGGGCTGAACCTGCAAAACCTGGGGCTCAGTGAAGATGTCCTGCAAG 69  
QY 73 GCTTCTGGCTACACCTTTTAACTACTGGATTCACTGGGTGAACAGAGGCTGGACAG 132  
DB 70 GCTTCTGGCTACACCTTTTAACTACTGGATTCACTGGGTGAACAGAGGCTGGACAG 129  
QY 133 GGTCTGAATGGATTGGATACATTAACTCTGCACCTGGTTCACCTTTTCAATCAGGAC 192  
DB 130 GGTCTGAATGGATTGGATATATTAACTCTGCACCTGGGTATCTGAATACATCAGAAG 189  
QY 193 TTTGAGGACAGGCGACCTTTGACCGCAGACAAAGTCTCCACACAGCCTACATGACGCTG 252  
DB 190 TTCAAGGACAGGCGACATTTGACTGCAGACAAATCTCCAGCAGACGCTACATGCAACTG 249  
QY 253 ACCAGCTGACATCTGAGGACTCTTCACTTATTACTGTGCAAGAGGGGTACGACGG 312  
DB 250 AGCAGCTGACATTTGAGGACTCTGCACTTATTACTGTGCAAGAGGGGGG---GGGTC 306  
QY 313 TTTGACTCTGGGGCAAGGACCACTCTCAGTCTCTCAAGTCTCTCA 354  
DB 307 TTTGACTCTGGGGCAAGGACCACTCTCAGTCTCTCAAGTCTCTCA 348

## RESULT 5

US-10-482-532-6  
; Sequence 6, Application US/10482532  
; Publication No. US20050063970A1

GENERAL INFORMATION:  
; APPLICANT: Reiter, Yoram  
; APPLICANT: Lev, Avital  
; TITLE OF INVENTION: METHODS AND PHARMACEUTICAL COMPOSITIONS FOR IMMUNE DECEPTION,  
; PARTICULARLY USEFUL IN THE TREATMENT OF CANCER  
; FILE REFERENCE: 71506-B-PCT-US  
; CURRENT APPLICATION NUMBER: US/10/482,532  
; PRIOR FILING DATE: 2003-12-19  
; PRIOR APPLICATION NUMBER: PCT/IL02/00478  
; PRIOR FILING DATE: 2002-06-18  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 348  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: atacvH sequence - a part of B2M-aTAc (dsFv)  
US-10-482-532-6

Query Match 75.4%; Score 266.8; DB 19; Length 348;  
Best Local Similarity 88.3%; Pred. No. 3.6e-78;  
Matches 302; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

QY 13 CTCGAGCAGCTGGGGCTGAACCTGCAAAACCTGGGGCTCAGTGAAGATGTCCTGCAAG 72  
DB 10 CTGACGAGTCTGGGGCTGAACCTGCAAAACCTGGGGCTCAGTGAAGATGTCCTGCAAG 69  
QY 73 GCTTCTGGCTACACCTTTTAACTACTGGATTCACTGGGTGAACAGAGGCTGGACAG 132  
DB 70 GCTTCTGGCTACACCTTTTAACTACTGGATTCACTGGGTGAACAGAGGCTGGACAG 129  
QY 133 GGTCTGAATGGATTGGATACATTAACTCTGCACCTGGTTCACCTTTTCAATCAGGAC 192  
DB 130 GGTCTGAATGGATTGGATATATTAACTCTAGCACTGGGTATCTGANTACATCAGAAG 189

QY 193 TTTGAGGACAGGCGACCTTTGACCGCAGACAAAGTCTCCACACAGCCTACATGACGCTG 252  
DB 190 TTCAAGGACAAAGGCGACATTTGACTGCAGACAAATCTCCAGCAGCCTACATGCAACTG 249  
QY 253 ACCAGCTGACATCTGAGGACTCTTCACTTATTACTGTGCAAGAGAGGGGTACGACGG 312  
DB 250 AGCAGCTGACATTTGAGGACTCTGCACTTATTACTGTGCAAGAGGGGGG---GGGTC 306  
QY 313 TTTGACTCTGGGGCAAGGACCACTCTCAGTCTCTCAAGTCTCTCA 354  
DB 307 TTTGACTCTGGGGCAAGGACCACTCTCAGTCTCTCAAGTCTCTCA 348

## RESULT 6

US-10-489-626-4  
; Sequence 4, Application US/10489626  
; Publication No. US20050079170A1

GENERAL INFORMATION:  
; APPLICANT: LeGALL, Fabrice  
; APPLICANT: KIPRIYANOV, Sergey  
; APPLICANT: REUSCH, Uwe  
; APPLICANT: MOLDENHAUER, Gerhard  
; APPLICANT: LITTLE, Melvyn  
; TITLE OF INVENTION: DIMERIC AND MULTIMERIC ANTIGEN BINDING STRUCTURE  
; FILE REFERENCE: 03528.0142.PCUS00  
; CURRENT APPLICATION NUMBER: US/10/489,626  
; CURRENT FILING DATE: 2004-03-12  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 4  
; LENGTH: 1817  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Plasmid  
US-10-489-626-4

Query Match 73.8%; Score 261.2; DB 19; Length 1817;  
Best Local Similarity 86.1%; Pred. No. 4.5e-76;  
Matches 303; Conservative 0; Mismatches 43; Indels 6; Gaps 1;

QY 9 GCTCTCGAGCAGCCTGGGGCTGAACTGGGCAAAACCTGGGGCTCAGTGAAGATGTCCTG 68  
DB 1031 GCAGCTGCAGCAGCTGGGGCTGAACTGGGCAAAACCTGGGGCTCAGTGAAGATGTCCTG 1090  
QY 69 CAAGGCTTCTGGCTACACCTTTTAACTACTGATTCACCTGGGTGAACAGAGGCTGG 128  
DB 1091 CAAGGCTTCTGGCTACACCTTTTAACTACTGATTCACCTGGGTGAACAGAGGCTGG 1150  
QY 129 ACAGGCTGAAATGGATTGGATATCAATTAATCTGCGCACTGGTTCACCTTTTACAATCA 188  
DB 1151 ACAGGCTGGAATGGATTGGATATCAATTAATCTGCGCACTGGTTCACCTTTTACAATCA 1210  
QY 189 GGACTTTTCAGGACAGGCGCACTTTGACCGCAGACAAAGTCTCCACACAGCCTACATCA 248  
DB 1211 GAAGTTCAAGGACAAAGGCGCAATTTGACTACAGACAAATCTCCAGCAGCCTACATCA 1270  
QY 249 GCTCAGCAGCCTGACATCTGAGGACTCTTCACTTACTTGTGCAAGA-----GAGGG 302  
DB 1271 ACTGAGCAGCTGACATCTGAGGACTCTGCACTTATTACTGTGCAAGATATTATGATGA 1330  
QY 303 GTACGAGGGTTTGAATCTCTGGGGCAAGGCAACCACTCTCAGTCTCTCA 354  
DB 1331 TCATTACAGCCTTGACTACTGGGGCAAGGCAACCACTCTCAGTCTCTCA 1382

## RESULT 7

US-10-489-626-11  
; Sequence 11, Application US/10489626  
; Publication No. US20050079170A1

GENERAL INFORMATION:  
; APPLICANT: LeGALL, Fabrice

APPLICANT: KIPRIYANOV, Sergey  
APPLICANT: REUSCH, Uwe  
APPLICANT: MOLDENHAUER, Gerhard  
APPLICANT: LITTLE, Melvyn  
TITLE OF INVENTION: DIMERIC AND MULTIMERIC ANTIGEN BINDING STRUCTURE  
FILE REFERENCE: 03528.0142.PCUS00  
CURRENT APPLICATION NUMBER: US/10/489,626  
CURRENT FILING DATE: 2004-03-12  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 11  
LENGTH: 6844  
TYPE: DNA  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: plasmid  
US-10-489-626-11

Query Match 73.8%; Score 261.2; DB 19; Length 6844;  
Best Local Similarity 86.1%; Pred. No. 6.9e-76;  
Matches 303; Conservative 0; Mismatches 43; Indels 6; Gaps 1;

Qy 9 GCTGCTCGAGCAGCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTG 68  
Db 2195 GCAGCTGAGCAGCTGGGGCTGAACCTGGCAAGACCTGGGGCTCAGTGAAGATGTCCTG 2254  
Qy 69 CAAGGCTTCTGGCTACACCTTTTACTTAACCTGAGTTCACCTGGGTGAAACAGAGGCTGG 128  
Db 2255 CAAGGCTTCTGGCTACACCTTTTACTAGGTACACGATGACCTGGGTAAACAGAGGCTGG 2314  
Qy 129 ACAGGGTCTGAATGGATGGATGATACATTAATCTTGGCCACTGGTCCACTTCTTACAATCA 188  
Db 2315 ACAGGGTCTGAATGGATGGATGATACATTAATCTTGGCCACTGGTCCACTTCTTACAATCA 2374  
Qy 189 GGACTTTCAGACAGGCGCCACTTTGACCGAGACAAGTCTCCACACAGCAGCTACATGCA 248  
Db 2375 GAAGTTCAAGACAAGGCGCCACTTTGACTACAGACAATCTCCACAGCAGCTACATGCA 2434  
Qy 249 GCTGACGAGCTGACATCTGAGGACTCTTCACTGATCTTACTTACTGTGCAAGA-----GAGGG 302  
Db 2435 ACTGAGCAGCTGACATCTGAGGACTCTGAGTCTTACTTACTGTGCAAGATATTATGATGA 2494  
Qy 303 GTAGCAGGGTTGACTCTCTGGGGCCAGGACACCTCTCACTCTCACTCTCTCA 354  
Db 2495 TCATTACAGCCTTGACTACTTGGGGCCAGGACCACTCTCACTCTCTCTCA 2546

RESULT 8  
US-10-682-845-43  
; Sequence 43, Application US/10682845  
; Publication No. US20040162411A1  
; GENERAL INFORMATION:  
; APPLICANT: Lanzavecchia, Antonio  
; TITLE OF INVENTION: Potent T cell modulating molecules  
; FILE REFERENCE: G2296 US  
; CURRENT APPLICATION NUMBER: US/10/682,845  
; PRIOR FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/419,149  
; PRIOR FILING DATE: 2002-10-18  
; PRIOR APPLICATION NUMBER: CA 2,403,313  
; PRIOR FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 43  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Variable region of OKT3 without signal peptide  
US-10-682-845-43

Query Match 73.5%; Score 260.2; DB 18; Length 357;  
Best Local Similarity 86.0%; Pred. No. 5.7e-76;

Matches 302; Conservative 0; Mismatches 43; Indels 6; Gaps 1;  
Qy 10 CTGCTCGAGCAGCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
Db 7 CAGCTGAGCAGTCTGGGGCTGAACCTGGCAAGACCTGGGGCTCAGTGAAGATGTCCTGC 66  
Qy 70 AAGGCTTCTGGCTACACCTTTTACTTAACCTGATGATCACTGGGGTGAACAGAGGCTTGA 129  
Db 67 AAGGCTTCTGGCTACACCTTTTACTTAGGTACAGATGCACTGGGTAAACAGAGGCTTGA 126  
Qy 130 CAGGCTCTGAATGGATGGATGATACATTAATCTCCACTGGTTCACCTTCTTACAATCAG 189  
Db 127 CAGGCTCTGAATGGATGGATGATACATTAATCTCCACTGGTTCACCTTCTTACAATCAG 186  
Qy 190 GACTTTCAGGACAGGCGCCACTTTGACCGCAGACAAGTCTCCACACAGCCTACATGCA 249  
Db 187 AAGTTCAAGGACAGGCGCCACTTTGACTACAGACAATCTCCACAGCAGCCTACATGCA 246  
Qy 250 CTGACGAGCTGACATCTGAGGACTCTTCACTGATCTTACTTGTGCAAGA-----GAGGG 303  
Db 247 CTGACGAGCTGACATCTGAGGACTCTGAGTCTTACTTGTGCAAGATATTATGATGAT 306  
Qy 304 TACGACGGTTGACTCTCTGGGGCCAGGACCACTCTCACTCTCACTCTCTCA 354  
Db 307 CATTACTGCTTGACTACTTGGGGCCAGGACCACTCTCACTCTCTCTCA 357

RESULT 9  
US-10-682-845-41  
; Sequence 41, Application US/10682845  
; Publication No. US20040162411A1  
; GENERAL INFORMATION:  
; APPLICANT: Lanzavecchia, Antonio  
; TITLE OF INVENTION: Potent T cell modulating molecules  
; FILE REFERENCE: G2296 US  
; CURRENT APPLICATION NUMBER: US/10/682,845  
; PRIOR FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/419,149  
; PRIOR FILING DATE: 2002-10-18  
; PRIOR APPLICATION NUMBER: CA 2,403,313  
; PRIOR FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 41  
; LENGTH: 417  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; OTHER INFORMATION: OKT3 heavy chain variable region with signal peptide  
US-10-682-845-41

Query Match 73.5%; Score 260.2; DB 18; Length 417;  
Best Local Similarity 86.0%; Pred. No. 6e-76;  
Matches 302; Conservative 0; Mismatches 43; Indels 6; Gaps 1;

Qy 10 CTGCTCGAGCAGCTGGGGCTGAACCTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69  
Db 67 CAGCTGAGCAGTCTGGGGCTGAACCTGGCAAGACCTGGGGCTCAGTGAAGATGTCCTGC 126  
Qy 70 AAGGCTTCTGGCTACACCTTTTACTTAACCTGATGATCACTGGGGTGAACAGAGGCTTGA 129  
Db 127 AAGGCTTCTGGCTACACCTTTTACTTAGGTACAGATGCACTGGGTAAACAGAGGCTTGA 186  
Qy 130 CAGGCTCTGAATGGATGGATGATACATTAATCTCCACTGGTTCACCTTCTTACAATCAG 189  
Db 187 CAGGCTCTGAATGGATGGATGATACATTAATCTCCACTGGTTCACCTTCTTACAATCAG 246  
Qy 190 GACTTTCAGGACAGGCGCCACTTTGACCGCAGACAAGTCTCCACACAGCCTACATGCA 249  
Db 247 AAGTTCAAGGACAGGCGCCACTTTGACTACAGACAATCTCCACAGCAGCCTACATGCA 306  
Qy 250 CTGACGAGCTGACATCTGAGGACTCTTCACTGATCTTACTTGTGCAAGA-----GAGGG 303

```
Db 307 CTGACGCGCTGACATCTGAGGACTCTGCGCTATTACTGTGCAAGATATTATGATGAT 366
Qy 304 TAGGACGGTTGACTCTCTGGGGCCAGGACCACTCTCAGAGTCTCTCA 354
Db 367 CATTACTGCTTGACTACTGCGGGCCAGGACCACTCTCAGAGTCTCTCA 417

RESULT 10
US-10-682-845-39
; Sequence 39, Application US/10682845
; Publication No. US20040162411A1
; GENERAL INFORMATION:
; APPLICANT: Lanzavecchia, Antonio
; TITLE OF INVENTION: Potent T cell modulating molecules
; FILE REFERENCE: G2296 US
; CURRENT APPLICATION NUMBER: US/10/682,845
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/419,149
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: CA 2,403,313
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 1570
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: OKT3 heavy chain with signal peptide
US-10-682-845-39

Query Match 73.5%; Score 260.2; DB 18; Length 1570;
Best Local Similarity 86.0%; Pred. No. 9.2e-76;
Matches 302; Conservative 0; Mismatches 43; Indels 6; Gaps 1;

Qy 10 CTGCTCGAGCAGCCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69
Db 104 CAGCTGAGCAGTCTGGGGCTGAACTGGCAAGACCTGGGGCTCAGTGAAGATGTCCTGC 163
Qy 70 AAGGCTTCTGCTACACCTTTTACTACTTACTTACTTACTTACTTACTTACTTACTTACTT 129
Db 164 AAGGCTTCTGCTACACCTTTTACTAGGTACAGATGCACTGGTAAACAGAGGCTGGA 223
Qy 130 CAGGCTCGAATGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 189
Db 224 CAGGCTCGAATGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 283
Qy 190 GACTTTCAGGACAGGCGCACTTTGACCGCAGACCAAGTCTCCACACAGCCTACATGCG 249
Db 284 AAGTTCAAGGACAGGCGCACTTTGACCGCAGACCAATCTCCAGCAGCCTACATGCGA 343
Qy 250 CTGACCGCTTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGA-----GAGGG 303
Db 344 CTGACCGCTTGACATCTGAGGACTCTTTCAGTCTATTACTGTGCAAGATATTATGATGAT 403
Qy 304 TAGGACGGTTGACTCTCTGGGGCCAGGACCACTCTCAGAGTCTCTCA 354
Db 404 CATTACTGCTTGACTACTGCGGGCCAGGACCACTCTCAGAGTCTCTCA 454

RESULT 11
US-10-059-261-109
; Sequence 109, Application US/10059261
; Publication No. US20030077826A1
; GENERAL INFORMATION:
; APPLICANT: EDELMAN, LENA
; APPLICANT: JACOTOT, ETIENNE DANIEL FRANCOIS
; APPLICANT: BRIAND, JEAN-PAUL
; TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
; TITLE OF INVENTION: SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
; TITLE OF INVENTION: FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
; FILE REFERENCE: 03495.0216
```

```
; CURRENT APPLICATION NUMBER: US/10/059,261
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/265,594
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 109
; LENGTH: 10511
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector pACgp67-ScFv350
; NAME/KEY: CDS
; LOCATION: (1)..(10509)
US-10-059-261-109

Query Match 73.5%; Score 260.2; DB 14; Length 10511;
Best Local Similarity 86.7%; Pred. No. 1.7e-75;
Matches 299; Conservative 0; Mismatches 43; Indels 3; Gaps 1;

Qy 13 CTCGAGCAGCCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGATGTCCTGCAAG 72
Db 4309 CTGACGAGTCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGCTGTCTCTGCAAG 4368
Qy 73 GCTTCTGGCTACACCTTTTACTAACTACTGGAATTCATCTGGGTGAAACAGAGGCTGGACG 132
Db 4369 GCTTCTGGCCACACCTTTTACTAGCTACTGGAATTCATCTGGGTGAAACAGAGGCTGGACG 4428
Qy 133 GGTCTGAAATGGATTGGATACATTAATCTCCACTGGTTCACCTTCTTACAAATCAGGAC 192
Db 4429 GGTCTGGATGGATTGGATACATTAATCTTAGCAGTGGTTATTTAAGTACAAATCAGGAC 4488
Qy 193 TTTCAAGGACAGGCGCACTTTGACCGCAGACCAAGTCTCCACACAGCCTACATGAGCTG 252
Db 4489 TTTCAAGGACAGGCGCACTTTGACCGCAGACCAAGTCTCCACACAGCCTACATGAGCTG 4548
Qy 253 ACCAGCTGACATCTGAGGACTCTTCACTTATTTACTGTGCAAGAGGGGTACACGGG 312
Db 4549 AGCAGCTGACATGATGAGGACTCTGCACTTATTACTGTGCAAGGCGAGCTCAGGCTAG 4608
Qy 313 ---TTTGACTCTCTGGGGCCAGGACCACTCTCAGAGTCTCTCA 354
Db 4609 ACCTTTGACTACTGGGGCCAGGACCACTCTCAGAGTCTCTCA 4653

RESULT 12
US-10-627-649-109
; Sequence 109, Application US/10627649
; Publication No. US20040265300A1
; GENERAL INFORMATION:
; APPLICANT: EDELMAN, LENA
; APPLICANT: JACOTOT, ETIENNE DANIEL FRANCOIS
; APPLICANT: BRIAND, JEAN-PAUL
; TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
; TITLE OF INVENTION: SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
; TITLE OF INVENTION: FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
; FILE REFERENCE: 02356-0083
; CURRENT APPLICATION NUMBER: US/10/627,649
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: 60/265,594
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 109
; LENGTH: 10511
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector pACgp67-ScFv350
```

```

; NAME/KEY: CDS
; LOCATION: (1)..(10509)
US-10-627-649-109

Query Match 73.5%; Score 260.2; DB 18; Length 10511;
Best Local Similarity 86.7%; Pred. No. 1.7e-75;
Matches 299; Conservative 0; Mismatches 43; Indels 3; Gaps 1;

Qy 13 CTGAGCAGCCTGGGCGTGAACCTGGCAAAACCTGGGGCCCTCAGTGAAAGATGTCCTGCAAG 72
Db 4309 CTGAGCAGCTCTGGGCGTGAACCTGGCAAAACCTGGGGCCCTCAGTGAAAGATGTCCTGCAAG 4368

Qy 73 GCTTCTGGCTACACCTTTTACTACTACTGATCTGATTTCACTGGGTGCAACAGAGGCGCTGCACAG 132
Db 4369 GCTTCTGGCCACACCTTTTACTAGCTACTGATGCTACTGGGTGCAACAGAGGCGCTGCACAG 4428

Qy 133 GGTCTGAAATGGATGGATACATTAATCTGCCACTGGTTCCACTTCTTACAATCAGGAC 192
Db 4429 GGTCTGAAATGGATGGATACATTAATCTTAGCAGTGGTTATATTAAAGTACAATCAGGAG 4488

Qy 193 TTTTCAGAGCAGCGCACTTTGACCGCAGACAAGTCTCTCACACAGCCTTACATGAGCTG 252
Db 4489 TTTTCAGAGCAGCGCACTTTGACCGCAGACAAGTCTCTCACACAGCCTTACATGAGCTG 4548

Qy 253 ACCAGCCTGACATCTGAGGACTCTTCAGTCTATTACTGTGCAAGAGAGGGGTACGACGGG 312
Db 4549 AGCAGCCTGACATATGAGGACTCTGAGTCTATTACTGTGCAAGGCGAGCTCAGGCTACG 4608

Qy 313 ---TTTGACTCTCGGGGCGCAAGGCACCACTCTCACAGTCTCTCTCA 354
Db 4609 ACCTTTGACTACTGGGGCGCAAGGCACCACTCTCACAGTCTCTCTCA 4653

RESULT 13
US-10-489-626-6
; Sequence 6, Application US/10489626
; Publication No. US20050079170A1
; GENERAL INFORMATION:
; APPLICANT: LEGALL, Fabrice
; APPLICANT: KIPRIYANOV, Sergey
; APPLICANT: REUSCH, Uwe
; APPLICANT: MOLDENHAUER, Gerhard
; APPLICANT: LITTLE, Melvyn
; TITLE OF INVENTION: DIMERIC AND MULTIMERIC ANTIGEN BINDING STRUCTURE
; FILE REFERENCE: 03528.0142.PCUS00
; CURRENT APPLICATION NUMBER: US/10/489,626
; CURRENT FILING DATE: 2004-03-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 1817
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Plasmid
US-10-489-626-6

Query Match 73.3%; Score 259.6; DB 19; Length 1817;
Best Local Similarity 85.8%; Pred. No. 1.5e-75;
Matches 302; Conservative 0; Mismatches 44; Indels 6; Gaps 1;

Qy 9 GCTGCTCGAGCAGCCTGGGGCTGAACTGCAAAAACCTGGGGCTCAGTGAAAGATGTCCTG 68
Db 194 GCAACTGCGACAGCTGCGGGCTGAACTGCAAAAACCTGGGGCTCAGTGAAAGATGTCCTG 253

Qy 69 CAAGGCTTCTGGCTACACCTTTTAACTACTGGAATTCAGTGGGTGAAACAGAGGCGCTGG 128
Db 254 CAAGGCTTCTGGCTACACCTTTTAACTACTGGAATTCAGTGGGTGAAACAGAGGCGCTGG 313

Qy 129 ACAGGCTGGAATGGAATGGATGATACATTAATCTCGCCTGCTGCTTCTTCAATCA 188
Db 314 ACAGGCTGGAATGGAATGGATGATACATTAATCTCGCCTGCTGCTTCTTCAATCA 373

; NAME/KEY: CDS
; LOCATION: (1)..(10509)
US-10-627-649-109

Query Match 73.5%; Score 260.2; DB 18; Length 10511;
Best Local Similarity 86.7%; Pred. No. 1.7e-75;
Matches 299; Conservative 0; Mismatches 43; Indels 3; Gaps 1;

Qy 13 CTGAGCAGCCTGGGCGTGAACCTGGCAAAACCTGGGGCCCTCAGTGAAAGATGTCCTGCAAG 72
Db 4309 CTGAGCAGCTCTGGGCGTGAACCTGGCAAAACCTGGGGCCCTCAGTGAAAGATGTCCTGCAAG 4368

Qy 73 GCTTCTGGCTACACCTTTTACTACTACTGATCTGATTTCACTGGGTGCAACAGAGGCGCTGCACAG 132
Db 4369 GCTTCTGGCCACACCTTTTACTAGCTACTGATGCTACTGGGTGCAACAGAGGCGCTGCACAG 4428

Qy 133 GGTCTGAAATGGATGGATACATTAATCTGCCACTGGTTCCACTTCTTACAATCAGGAC 192
Db 4429 GGTCTGAAATGGATGGATACATTAATCTTAGCAGTGGTTATATTAAAGTACAATCAGGAG 4488

Qy 193 TTTTCAGAGCAGCGCACTTTGACCGCAGACAAGTCTCTCACACAGCCTTACATGAGCTG 252
Db 4489 TTTTCAGAGCAGCGCACTTTGACCGCAGACAAGTCTCTCACACAGCCTTACATGAGCTG 4548

Qy 253 ACCAGCCTGACATCTGAGGACTCTTCAGTCTATTACTGTGCAAGAGAGGGGTACGACGGG 312
Db 4549 AGCAGCCTGACATATGAGGACTCTGAGTCTATTACTGTGCAAGGCGAGCTCAGGCTACG 4608

Qy 313 ---TTTGACTCTCGGGGCGCAAGGCACCACTCTCACAGTCTCTCTCA 354
Db 4609 ACCTTTGACTACTGGGGCGCAAGGCACCACTCTCACAGTCTCTCTCA 4653

RESULT 13
US-10-489-626-6
; Sequence 6, Application US/10489626
; Publication No. US20050079170A1
; GENERAL INFORMATION:
; APPLICANT: LEGALL, Fabrice
; APPLICANT: KIPRIYANOV, Sergey
; APPLICANT: REUSCH, Uwe
; APPLICANT: MOLDENHAUER, Gerhard
; APPLICANT: LITTLE, Melvyn
; TITLE OF INVENTION: DIMERIC AND MULTIMERIC ANTIGEN BINDING STRUCTURE
; FILE REFERENCE: 03528.0142.PCUS00
; CURRENT APPLICATION NUMBER: US/10/489,626
; CURRENT FILING DATE: 2004-03-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 1817
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Plasmid
US-10-489-626-6

Query Match 73.3%; Score 259.6; DB 19; Length 1817;
Best Local Similarity 85.8%; Pred. No. 1.5e-75;
Matches 302; Conservative 0; Mismatches 44; Indels 6; Gaps 1;

Qy 9 GCTGCTCGAGCAGCCTGGGGCTGAACTGCAAAAACCTGGGGCTCAGTGAAAGATGTCCTG 68
Db 194 GCAACTGCGACAGCTGCGGGCTGAACTGCAAAAACCTGGGGCTCAGTGAAAGATGTCCTG 253

Qy 69 CAAGGCTTCTGGCTACACCTTTTAACTACTGGAATTCAGTGGGTGAAACAGAGGCGCTGG 128
Db 254 CAAGGCTTCTGGCTACACCTTTTAACTACTGGAATTCAGTGGGTGAAACAGAGGCGCTGG 313

Qy 129 ACAGGCTGGAATGGAATGGATGATACATTAATCTCGCCTGCTGCTTCTTCAATCA 188
Db 314 ACAGGCTGGAATGGAATGGATGATACATTAATCTCGCCTGCTGCTTCTTCAATCA 373

; NAME/KEY: CDS
; LOCATION: (1)..(10509)
US-10-627-649-109

Query Match 73.5%; Score 260.2; DB 18; Length 10511;
Best Local Similarity 86.7%; Pred. No. 1.7e-75;
Matches 299; Conservative 0; Mismatches 43; Indels 3; Gaps 1;

Qy 13 CTGAGCAGCCTGGGCGTGAACCTGGCAAAACCTGGGGCCCTCAGTGAAAGATGTCCTGCAAG 72
Db 4309 CTGAGCAGCTCTGGGCGTGAACCTGGCAAAACCTGGGGCCCTCAGTGAAAGATGTCCTGCAAG 4368

Qy 73 GCTTCTGGCTACACCTTTTACTACTACTGATCTGATTTCACTGGGTGCAACAGAGGCGCTGCACAG 132
Db 4369 GCTTCTGGCCACACCTTTTACTAGCTACTGATGCTACTGGGTGCAACAGAGGCGCTGCACAG 4428

Qy 133 GGTCTGAAATGGATGGATACATTAATCTGCCACTGGTTCCACTTCTTACAATCAGGAC 192
Db 4429 GGTCTGAAATGGATGGATACATTAATCTTAGCAGTGGTTATATTAAAGTACAATCAGGAG 4488

Qy 193 TTTTCAGAGCAGCGCACTTTGACCGCAGACAAGTCTCTCACACAGCCTTACATGAGCTG 252
Db 4489 TTTTCAGAGCAGCGCACTTTGACCGCAGACAAGTCTCTCACACAGCCTTACATGAGCTG 4548

Qy 253 ACCAGCCTGACATCTGAGGACTCTTCAGTCTATTACTGTGCAAGAGAGGGGTACGACGGG 312
Db 4549 AGCAGCCTGACATATGAGGACTCTGAGTCTATTACTGTGCAAGGCGAGCTCAGGCTACG 4608

Qy 313 ---TTTGACTCTCGGGGCGCAAGGCACCACTCTCACAGTCTCTCTCA 354
Db 4609 ACCTTTGACTACTGGGGCGCAAGGCACCACTCTCACAGTCTCTCTCA 4653

RESULT 13
US-10-489-626-6
; Sequence 6, Application US/10489626
; Publication No. US20050079170A1
; GENERAL INFORMATION:
; APPLICANT: LEGALL, Fabrice
; APPLICANT: KIPRIYANOV, Sergey
; APPLICANT: REUSCH, Uwe
; APPLICANT: MOLDENHAUER, Gerhard
; APPLICANT: LITTLE, Melvyn
; TITLE OF INVENTION: DIMERIC AND MULTIMERIC ANTIGEN BINDING STRUCTURE
; FILE REFERENCE: 03528.0142.PCUS00
; CURRENT APPLICATION NUMBER: US/10/489,626
; CURRENT FILING DATE: 2004-03-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 1817
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Plasmid
US-10-489-626-6

Query Match 73.3%; Score 259.6; DB 19; Length 1817;
Best Local Similarity 85.8%; Pred. No. 1.5e-75;
Matches 302; Conservative 0; Mismatches 44; Indels 6; Gaps 1;

Qy 9 GCTGCTCGAGCAGCCTGGGGCTGAACTGCAAAAACCTGGGGCTCAGTGAAAGATGTCCTG 68
Db 194 GCAACTGCGACAGCTGCGGGCTGAACTGCAAAAACCTGGGGCTCAGTGAAAGATGTCCTG 253

Qy 69 CAAGGCTTCTGGCTACACCTTTTAACTACTGGAATTCAGTGGGTGAAACAGAGGCGCTGG 128
Db 254 CAAGGCTTCTGGCTACACCTTTTAACTACTGGAATTCAGTGGGTGAAACAGAGGCGCTGG 313

Qy 129 ACAGGCTGGAATGGAATGGATGATACATTAATCTCGCCTGCTGCTTCTTCAATCA 188
Db 314 ACAGGCTGGAATGGAATGGATGATACATTAATCTCGCCTGCTGCTTCTTCAATCA 373

; NAME/KEY: CDS
; LOCATION: (1)..(10509)
US-10-627-649-109

Query Match 73.5%; Score 260.2; DB 18; Length 10511;
Best Local Similarity 86.7%; Pred. No. 1.7e-75;
Matches 299; Conservative 0; Mismatches 43; Indels 3; Gaps 1;

Qy 13 CTGAGCAGCCTGGGCGTGAACCTGGCAAAACCTGGGGCCCTCAGTGAAAGATGTCCTGCAAG 72
Db 4309 CTGAGCAGCTCTGGGCGTGAACCTGGCAAAACCTGGGGCCCTCAGTGAAAGATGTCCTGCAAG 4368

Qy 73 GCTTCTGGCTACACCTTTTACTACTACTGATCTGATTTCACTGGGTGCAACAGAGGCGCTGCACAG 132
Db 4369 GCTTCTGGCCACACCTTTTACTAGCTACTGATGCTACTGGGTGCAACAGAGGCGCTGCACAG 4428

Qy 1
```

Db 164 AAGGCTTCGGCTACACCTTTACTAGGTACACGATGCACTGGGTAAACAGAGGCTTGA 223  
Qy 130 CAGGCTCTGAATGATGGATGATACATTAATCTCTGCACTGTTCCACTTCTTCAATCAG 189  
Db 224 CAGGCTCTGGAATGATGGATGATACATTAATCTAGCCGTGGTTACTTAATTAATCAG 283  
Qy 190 GACTTTCAGGACAGGCGCACTTTGACCGCAGACAAAGTCTCCACACAGCCTACATCAG 249  
Db 284 AAGTTCAAGGACAAGGCCACATTAATGACTACAGACAAATCTCCAGCAGCCTACATGCA 343  
Qy 250 CTGACACGCTGACATCTGAGGACTTTCAGTCTTACTTACTTGTGCAAGA-----GAGGGG 303  
Db 344 CTGACAGCCTGACATCTGAGGACTCTGCACTCTTACTTGTGCAAGATATTATGATGAT 403  
Qy 304 TACGACGGGTTGACTCTCTGGGGCAAGGCAACCACTCTCACAGTCTCTCA 354  
Db 404 CATTAAGTCTGCTGACTACTGGGGCAAGGCCACCACTCTCACAGTCTCTCA 454

## RESULT 15

US-10-704-352-6  
; Sequence 6, Application US/10704352  
; Publication No. US20040071693A1  
; GENERAL INFORMATION:  
; APPLICANT: Adair, John R.  
; Achwal, Diljeet S.  
; Entage, John S.  
; TITLE OF INVENTION: Humanised Antibodies  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/704,352  
; FILING DATE: 07-Nov-2003  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/795,515  
; FILING DATE: 28-FEB-2001  
; APPLICATION NUMBER: 08/846,658  
; FILING DATE: 01-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Trujillo, Doreen Yatko  
; REGISTRATION NUMBER: 35,719  
; REFERENCE/DOCKET NUMBER: CARP-0057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1570 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 41..1444  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 98..1444  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-10-704-352-6

Query Match 73.1%; Score 258.6; DB 17; Length 1570;  
Best Local Similarity 85.8%; Pred. No. 3.1e-75;  
Matches 301; Conservative 0; Mismatches 44; Indels 6; Gaps 1;  
Qy 10 CTGCTCGAGCAGCCTGGGGCTGAACTGGCAAAACCTGGGGCTCAGTGAAGATGTCTCTGC 69  
Db 104 CAGCTGCAAGCAGTCTGGGGCTGAACTGGCAAGACCTGGGGCTCAGTGAAGATGTCTCTGC 163  
Qy 70 AAGCTTCTGGCTACACCTTTTACTTAACCTACTGGAATCACTGGGGTGAACAGAGGCTTGA 129  
Db 164 AAGGCTTCTGGCTACACCTTTTACTTAGGTACACGATGCACTGGGTAAACAGAGGCTTGA 223  
Qy 130 CAGGCTCTGAATGATGGATGATACATTAATCTCTGCACTGTTCCACTTCTTCAATCAG 189  
Db 224 CAGGCTCTGAATGATGGATGATACATTAATCTTAGCCGTGGTTACTTAATTAATCAG 283  
Qy 190 GACTTTCAGACAGGCGCACTTTGACCGCAGACAAAGTCTCCACACAGCCTACATGCA 249  
Db 284 AAGTTCAAGGACAAGGCCACATTAATGACTACAGACAAATCTCCAGCAGCCTACATGCA 343  
Qy 250 CTGACACGCTGACATCTGAGGACTTTCAGTCTTACTTACTTGTGCAAGA-----GAGGGG 303  
Db 344 CTGACAGCCTGACATCTGAGGACTCTGCACTCTTACTTGTGCAAGATATTATGATGAT 403  
Qy 304 TACGACGGGTTGACTCTCTGGGGCAAGGCAACCACTCTCACAGTCTCTCA 354  
Db 404 CATTAAGTCTGCTGACTACTGGGGCAAGGCCACCACTCTCACAGTCTCTCA 454  
Search completed: June 4, 2005, 13:43:36  
Job time : 856.914 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:13:21 ; Search time 5343.6 Seconds  
(without alignments)  
2521.660 Million cell updates/sec

Title: US-10-089-452-1

Perfect score: 354

Sequence: 1 gaggtcgagctgctcgagca.....ccactctcacagtctctca 354

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_hcc.\*

4: gb\_est3.\*

5: gb\_est4.\*

6: gb\_est5.\*

7: gb\_est6.\*

8: gb\_gsl1.\*

9: gb\_gsl2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	253.8	71.7	919	5	BQ938998 AGENCOURT
2	246.4	69.6	540	6	CA579750 K0741H07-
3	244	68.9	406	2	BF016722 uy34h12.y
4	243.2	68.7	473	6	CA580167 K0747G08-
5	243.2	68.7	519	6	CA580087 K0746G01-
6	241	68.1	724	4	BG962137 602826902
7	234.6	66.3	849	2	BF582051 602099230
8	231	65.3	729	4	BF100304 602885769
9	228.2	64.5	698	2	BF584024 602096264
10	227	64.1	529	6	CA577954 K0716F08-
11	226.6	64.0	847	2	BF165486 601773393
12	225	63.6	684	2	BF577847 602092155
13	225	63.6	920	4	BF647308 603279436
14	223.6	63.2	501	6	CA5707443 K0709F01-
15	223.6	63.2	501	6	CA579132 K0733B12-
16	223.6	63.2	501	6	CA579461 K0737G06-
17	223.6	63.2	531	6	CA577710 K0713C04-
18	223.6	63.2	702	4	BG966479 602832989
19	223.6	63.2	857	5	BU522453 AGENCOURT
20	223.4	63.1	771	2	BF140551 601787584
21	223.4	63.1	1039	5	BQ959855 AGENCOURT
22	223.2	63.1	637	6	BF724721 BY724721
23	222	62.7	503	6	CA577465 K0710A12-
24	222	62.7	506	6	CA577320 K0707H01-

25	222	62.7	507	6	CA579179	CA579179 K0733G11-
26	222	62.7	723	4	BG969577	BG969577 602837396
27	221.8	62.7	358	2	BE850886	BE850886 uw89H05.y
28	221.8	62.7	1576	3	AK007918	AK007918 Mus muscu
29	221.6	62.6	739	2	BE284158	BE284158 601099428
30	220.4	62.3	508	6	CA578554	CA578554 K0725E05-
31	220.2	62.2	724	2	BF168514	BF168514 601775412
32	218.8	61.8	667	6	BY724790	BY724790 BY724790
33	218.6	61.8	538	6	CA578968	CA578968 K0731A04-
34	218.6	61.8	769	2	BF168856	BF168856 601775314
35	218.6	61.8	1012	2	BF579009	BF579009 602096127
36	217	61.3	789	4	BI249422	BI249422 602995753
37	217	61.3	819	4	BG966397	BG966397 602832896
38	216.8	61.2	745	4	BG871607	BG871607 602790090
39	215.4	60.8	794	4	BI150371	BI150371 602915205
40	215.4	60.8	865	4	BG964435	BG964435 602832080
41	215.4	60.8	913	4	BG962314	BG962314 602827112
42	214	60.5	903	4	BI415690	BI415690 602988758
43	213.8	60.4	690	4	BG968682	BG968682 602836513
44	213.8	60.4	767	7	CO565860	CO565860 AGENCOURT
45	212.6	60.1	793	2	BF136093	BF136093 601783539

## ALIGNMENTS

RESULT 1  
BQ938998  
LOCUS  
DEFINITION BQ938998 919 bp mRNA linear EST 21-AUG-2002  
IMAGE: 6478021 5', mRNA sequence.  
ACCESSION BQ938998.1 GI:22354476  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE  
1 (bases 1 to 919)  
NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAMA4020 row: 0 column: 14  
High quality sequence stop: 678.

## FEATURES

source  
1..919  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clones="IMAGE:6478021"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Co24"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."

## ORIGIN

Query Match 71.7%; Score 253.8; DB 5; Length 919;  
Best Local Similarity 83.5%; Pred. No. 8e-66;  
Matches 288; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 10 CTGCTCAGCAGCTGGGGCTGAAGTGGGCTCAAAACCTGGGGCTCAGTGAAGATGTCCTGC 69

```
Db 96 CAGCTTCAGCAGTCTGGGCGCTGAACCTGCAAAACCTGGAGCCTCAGTGAAGATGCTCTGC 155
Qy 70 AAGGCTTCGCTGCTACACCTTTACTAACTACTGGAATTCACCTGGGTGAACACAGAGCCCTGGA 129
Db 156 AAGGCTTCGCTGCTACACCTTTACTAACTACTGGAATTCACCTGGGTGAACACAGAGCCCTGGA 215
Qy 130 CAGGCTTCGAATGGAATGGATGATACATTAATCCTGCCACTGGTTCCACTTCTTACAATCAG 189
Db 216 CAGGCTTCGGAATGGAATGGATGATACATTAATCCTGCCACTGGTTCCACTTCTTACAATCAG 275
Qy 190 GACTTTCAGGACAGGCGCACTTTGACCGCAGACAGTCTCCACACAGCCTCATGCGAG 249
Db 276 AAGTTCAGGACAGGCGCACTTTGACCGCAGACAGTCTCCACACAGCCTCATGCGAG 335
Qy 250 CTGACACAGCCTGACATCTGAGGACTCTTCACTGCTATTACTGTCGAAGAGAGGGGTACGAC 309
Db 336 CTGACACAGCCTGACATCTGAGGACTCTTCACTGCTATTACTGTCGAAGAGAGGGGTACGAC 395
Qy 310 GGGTTTCACTCTCGGGGCCAAGGACCACTCTCAGCTCTCTCA 354
Db 396 AGGTTTGCTTACTGGGCGCAAGGAGCTCTGGTCACTGTCTCTGCA 440
```

```
RESULT 2
CA579750
LOCUS
DEFINITION K0741H07-5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)
cDNA Library (Long) Mus musculus cDNA clone NIA:K0741H07
IMAGE:30077562 5', mRNA sequence.
ACCESSION CA579750
VERSION CA579750.1 GI:25128141
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 540)
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
Alba, K., Taub, D., Longo, D.L., Keller, J. and Ko, M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)
JOURNAL Unpublished (2001)
COMMENT Other ESTs: K0741H07-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: K0741 row: H column: 07
Seq primer: M13 Reverse
High quality sequence stop: 540
POLYA=No.
```

```
FEATURES
source Location/Qualifiers
1..540
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6Ncr"
/db_xref="nigEST:K0741H07-5N"
/db_xref="taxon:10090"
/clones="NIA:K0741H07 IMAGE:30077562"
/tissue_type="Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-)"
/dev_stage="Age approx.10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)"
/notes="vector: pSPORT1 (Invitrogen); Site_1: Sali; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001)). [PMID: 11544199]". Total RNAs were
```

obtained from Drs. Dennis Taub, Dan Longo (National Institute on Aging, USA), Jonathan Keller (National Cancer Institute, USA). Double-stranded cDNAs were synthesized with an Oligo(dT) primer (Invitrogen): 5'-pGACTAGTCTAGATCGAGCGCCCTTTT-TTTT-3' from 0.9 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lona-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with Sali and NotI enzymes and cloned into Sali/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.1 kb. The library was constructed by Yulan Piao (NIA)."

## ORIGIN

```
Query Match 69.6%; Score 246.4; DB 6; Length 540;
Best Local Similarity 83.1%; Pred. No. 1.2e-63;
Matches 299; Conservative 0; Mismatches 46; Indels 15; Gaps 1;

Qy 10 CTGCTCAGCAGCGCTGGGCTGAACCTGGCAAAACCTGGGCGCTCAGTGAAGATGCTCTGC 69
Db 101 CAGCTGCAGCAGTCTGGGCTGAACCTGGCAAAACCTGGGCGCTCAGTGAAGATGCTCTGC 160
Qy 70 AAGGCTTCGGCTACACCTTTACTAACTACTGGAATTCACCTGGGTGAACACAGAGCCCTGGA 129
Db 161 AAGGCTTCGGCTACACCTTTACTAGTACTGGAATGCACTGGGTGAACACAGAGCCCTGGA 220
Qy 130 CAGGCTCTGAATGGATTGGATACATTAATCCTGCCACTGTTCCTCACTTCTTACAATCAG 189
Db 221 CAGGCTCTGAATGGATTGGATACATTAATCCTAGCAGTGGTTATCTAATGATCAATCAG 280
Qy 190 GACTTTCAGCAGCGGCCACTTTGACCGCAGACAGTCTCCACACAGCCTCATGCGAG 249
Db 281 AAGTTCAGGACAGGCGCACATGACTGTCAGACAGAAATCTCCAGCAGCAGCCTCATGCGAG 340
Qy 250 CTGACACAGCCTGACATCTGAGGACTCTTCACTGCTATTACTGTGCAAGAGGGGGCAT 301
Db 341 CTGAGCAGCCTGACATATGAGGACTCTGCACTGCTATTACTGTGCAAGAGGGGGCAT 400
Qy 302 -----GGTACGACGGGTTTGACTCTCTGGGGCGCAAGGCACCACTCTCAGAGTCTCTCTCA 354
Db 401 CTACTATGGTCTATGCTATGGACTACTGGGTCAAGGACCTCAGTCACCGTCTCTCTCA 460
```

## RESULT 3

```
BF016722
LOCUS
DEFINITION BF016722 406 bp mRNA linear EST 29-DEC-2000
similar to SW:FW02_MOUSE P01746 IG HEAVY CHAIN V REGION 93G7
PRECUSOR. ; mRNA sequence.
ACCESSION BF016722
VERSION BF016722.1 GI:10748054
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 406)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-i@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Washington University Genome Sequencing Center
```

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov/image/html/iresources.shtml](http://image.llnl.gov/image/html/iresources.shtml)

MGI:1422279  
Seq primer: -40RP from Gibco.

## FEATURES

```

Location/Qualifiers
1. 406
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:3661511"
/tissue_type="tumor, metastatic"
/lab_host="DH10B"
/clone_1ib="NCI CGAP Lu31"

```

/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

## ORIGIN

```
Query Match      68.9%; Score 244; DB 2; Length 406;
Best Local Similarity 88.3%; Pred. No. 6.2e-63;
Matches 265; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

Qy	10	CTGCTGAGCAGCCTGGGGCTGAACTGGGCAAAACCTGGGGCCTCAGTGAAGATGTCTCTGC	69
Db	99	 CAGCTCAGCAGTCTGGGGCTGAACCTGGCAAAACCTGGGGCCTCAGTGAAGATGTCTCTGC	158
Qy	70	AAGGCTTCTGGCTACACCTTTTACTAACTACTGGAATTCACCTGGGTGAAACAGAGGCCTGGA	129
Db	159	AAGGCTTCTGGCTACACCTTTTACTAGCTACTGATGCACTGGGTAAACAGAGGCCTGGA	218
Qy	130	CAGGGTCTGAATGGATTTGGATACATTAATCTTGCCCATCTGGTTCCACTCTTTACAATCAG	189
Db	219	CAGGGTCTGAATGGATTTGGATACATTAATCTTAGCAGTGGTTATATCTGAGTACAATCAG	278
Qy	190	GACTTTTCAGACAGGCGCACTTTGACCGCAGACAAGTCCTCCACACAGCCTCATGTCAG	249
Db	279	AAGTTTCAGACAGGCGCACTTTGATCTGCAGACAAATCTCTCAGACAGCCTCATGTCAG	338
Qy	250	CTGACCAAGCCTGACATCTGAGGACTCTTCAGTCTATTATCTGTGCAAGAGAGGGGTACGAC	309
Db	339	CTGACCAAGCCTGACATCTGAGGACTCTTCAGTCTATTATCTGTGCAAGATCGTTTTATTAC	398

## RESULT 4

CA580167	473 bp	linear	EST 19-NOV-2002
LOCUS			
DEFINITION			
K0747G08-5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-) CDNA Library (Long) Mus musculus cDNA clone NIA:K0747G08 IMAGE:30078127 5', mRNA sequence.			

ACCESSION  
CA580167

VERSION CA580167.1 GI:25128558

**KEYWORDS** EST.

**SOURCE** *Mus musculus* (house mouse)

ORGANISM MUS MUSCULUS

**Eukaryota; Metazoa; Chorda**

Mammalia; Eutheria; Rodent

REFERENCE: 1 (bases 1 to 473)

**AUTHORS** Piao, Y., Kargul, G.J., Dude

Aiba, K., Taub, D., Longo, D.

# TITLE Systematic Analyses of NIA

(Lin-/c-Kit-/Sca-1-) cDNA

JOURNAL Unpublished (2001)

COMMENT Other\_ESTs: K0747G08-3

Contact: Dawood B. Dudekul

Laboratory of Genetics

National Institute on Aging

333 Cassell Drive, Suite 4

Email: [cdna@lgsun.grc.nia](mailto:cdna@lgsun.grc.nia).

Plate: K0747 row: G column: 08  
Seq primer: M13 Reverse  
High quality sequence stop: 473  
POLYA=No.

**FEATURES**  
**source**

```

1 473
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="C57BL/6NCr"
/strain="57BL/6NCr"
/db_xref="naiest:K0747608-5N"
/db_xref="taxon:10090"
/clones="NTA:K0747608 IMAGE:30078127"
/tissue_types="Hematopoietic Stem Cell"
/dev_stage="G1-Sca-1")
/lab_host="Dh108"

```

```

/clone_lib="NTA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) cDNA Library (long)"
/note="vector: pSPORI (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIA (http://lgscn.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Drs. Dennis Taub, Dan Koller (National
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an Oligo(dT) primer [Invitrogen:

```

5'-pGACTAGTTCAGATCGGAGCGGGCGCGCTTTTTTTTTTT-3'] from purified by ethanol-precipitation. The cDNAs were ligated to Loxe-linker Lu-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORI plasmid vector. The DH10B *E. coli* host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.1 kb. The library was constructed by Yulan Piao (NIA)."

## ORIGIN

Query Match 68.7%; Score 243.2; DB 6; Length 473;  
Best Local Similarity 82.5%; Pred. No. 1.1e-62;  
Matches 297; Conservative 0; Mismatches 48; Indels 15; Gaps 1;

Qy	10	CTGCTCGAGCAGCCTGGGGCTGAACTTGGGAAACCTGGGGCCTCAGTGAAGATGTCTCTGC	69
Db	101	CAGCTCGAGCAGTCTCGGGCTGAACCTGGGAAAACCTGGGGCCTCAGTGAAGCTGTCTCTGC	160
Qy	70	AAGGCTTCTGGCTACACCTTTACTTACTACTGATTTCACTGGGTGAAACAGAGGCCTGGA	129
Db	161	AAGGCTTCTGGCTACACCTTCTACTAGCTACTTGGATGCACTCTGGGTAAACAGAGGCCTGGA	220
Qy	130	CAGGCTCTGAATGGATTTGGATACATTAATCTTCGCCACTGGTTTCCACTTCTTACAATCAG	189
Db	221	CAGGGTCTGGAATGGATTTGGATACGTTAATCTTAGCAGTGGTTATTAAGTACAATCAG	280
Qy	190	GACTTTTCAGGACAGGGGCCACTTTTGACCGCAGACAAGTCTCCACACAGGCCTACATGCAG	249
Db	281	AAGTTCAAGGACAGGGCCCAATTGACTGCGNACAATACTCTCAGACAAGCCTACATGCAG	340
Qy	250	CTGACCAAGCCTGACATCTCAGGACCTTTCAGTCTATTACTGTGCAAGAGAGG-----	301
Db	341	CTGAGCAGCCTGACATATGAGGACTCTGCAGTCTATTACTGTGCAAGAGGGAGGGGCAT	400
Qy	302	-----GGTACGACGGGTTTGACTCTCTGGGGCCAGGCACCACTCTCAAGTCTCTCTCA	354
Db	401	CTACTATGGGTCTATGCTATGGACTACTGGGGTCAAGGAACCTTCAGTCAACCGTCTCTCTCA	460

```

RESULT 5
CA580087      519 bp      mRNA      linear      EST 19-NOV-2002
LOCUS         K0746G01-5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)
DEFINITION    IMAGE:30078024 5', mRNA sequence.
ACCESSION     CA580087
VERSION       CA580087.1 GI:25128478
KEYWORDS      Mus musculus (house mouse)
SOURCE        Mus musculus
ORGANISM      Mus musculus
REFERENCE     1. (Bases 1 to 519)
AUTHORS       Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
              Aiba, K., Taub, D., Longo, D.L., Keller, J. and Ko, M.S.H.
              Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
              (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)
              Unpublished (2001)
JOURNAL       Other ESTs: K0746G01-3
COMMENT       Contact: Dawood B. Dudekula
              Laboratory of Genetics
              National Institute on Aging/National Institutes of Health
              333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
              Email: cdna@igsun.grc.nia.nih.gov
              Plate: K0746 row: G column: 01
              Seq primer: M13 Reverse
              High quality sequence stop: 519
              POLYA=No.
FEATURES      Location/Qualifiers
source        1..519
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6NCr"
              /db_xref="niaEST:K0746G01-5N"
              /db_xref="taxon:10090"
              /clone="NIA:K0746G01 IMAGE:30078024"
              /tissue_type="Hematopoietic Stem Cell
              (Lin-/c-Kit-/Sca-1-)"
              /dev_stage="Age approx.10 weeks old"
              /lab_host="DH10B"
              /clone_lib="NIA Mouse Hematopoietic Stem Cell
              (Lin-/c-Kit-/Sca-1-) cDNA Library (Long)"
              /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
              NotI; Mouse cDNA project by the Laboratory of Genetics,
              National Institute on Aging (NIA), Intramural Research
              program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
              a long-transcript enriched cDNA library (Ref. Genome Res.
              11: 1553-1558 (2001). [PMID: 1154199]). Total RNAs were
              obtained from Drs. Dennis Taub, Dan Longo (National
              Institute on Aging, USA), Jonathan Keller (National Cancer
              Institute, USA). Double-stranded cDNAs were synthesized
              with an Oligo(dT) primer [Invitrogen].
              5'-pGACTAGTTCTAGATCGGAGCGGCCCTTTT-3'] from
              0.9 ug of total RNA, treated with T4 DNA polymerase, and
              purified by ethanol-precipitation. The cDNAs were ligated
              to lone-linker L1-Sal4, purified by phenol/chloroform, and
              separated from free linkers by Centricon 100. Then, the
              cDNAs were amplified by long-range high fidelity PCR using
              Ex Taq polymerase (Takara) with a primer Sal4-S. The
              products were purified by phenol/chloroform and Centricon
              100. The cDNAs were digested with SalI and NotI enzymes
              and cloned into SalI/NotI site of pSPORT1 plasmid vector.
              The DH10B E. coli host was transformed with the ligation
              mixture by the standard chemical method. The average
              insert size is about 2.1 kb. The library was constructed
              by Yulan Fiao (NIA)."
ORIGIN
Query Match      68.7%; Score 243.2; DB 6; Length 519;
Best Local Similarity 82.5%; Pred. No. 1.2e-62;
Matches 297; Conservative 0; Mismatches 48; Indels 15; Gaps 1;

```

```

Qy 10 CTGCTCAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCCTCAGTGAAGATGTCCTGC 69
    |||
Db 101 CAGCTGCAGCAGTCTGGGGCTGAACCTGGCAAAACCTGGGGCCTCAGTGAAGATGTCCTGC 160
    |||
Qy 70 AAGGCTTCTGGCTACACCTTTTACTAACTACTGGGATTCACCTGGGTGAACAGAGCCTCGA 129
    |||
Db 161 AAGGCTTCTGGCTACACCTTCACTAGTACTGGATGCATGGGTGAAACAGAGCCTCGA 220
    |||
Qy 130 CAGGCTCTGAATGGATTGGATACATTAATCTCCCACTGGTTCCTTCTTACAATCAG 189
    |||
Db 221 CAGGCTCTGAATGGATTGGATAGTTAATCTTAGCAGTGGTTATATAGTACAATCAG 280
    |||
Qy 190 GACTTTGAGCAGAGGGCACTTTGACCGCAGACAAGTCTCCACACAGCCTCATGAGCAG 249
    |||
Db 281 AAGTTCGAAGGACAAGGCCACATTCAGTGCAGACAAATCTCCAGCAGACAGCCTCATGAGCAG 340
    |||
Qy 250 CTGACCAAGCTGACATCTGAGGACTCTTCTAGTCTTACTTCTGTGCAAGAGAGG----- 301
    |||
Db 341 CTGAGCAGCCTGCATATGAGGACTCTGCACTCTTACTTGTGCAAGAGAGGAGGGGCAT 400
    |||
Qy 302 -----GCTACGACGGGTTTGACTCTCTGGGGCCAGGCACCACTCTCACAGTCTCTCTCA 354
    |||
Db 401 CTACTATGGGCTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCTCA 460
    |||
RESULT 6
BG962137      724 bp      mRNA      linear      EST 12-JUN-2001
LOCUS         BG962137
DEFINITION    60286902F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4981585 5',
              mRNA sequence.
ACCESSION     BG962137
VERSION       BG962137.1 GI:14349774
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE     1. (Bases 1 to 724)
AUTHORS       NIH-MGC http://mgc.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapsb@mail.nih.gov
              Tissue Procurement: Jeffrey E. Green, M.D.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1AM10983 row: d column: 02
              High quality sequence stop: 713.
FEATURES      Location/Qualifiers
source        1..724
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="FVB/N"
              /db_xref="taxon:10090"
              /clone="IMAGE:4981585"
              /lab_host="DH10B (TI phage-resistant)"
              /clone_lib="NCI CGAP Co24"
              /note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
              Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
              Average insert size 1.6 kb. Constructed by Life
              Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Query Match      68.1%; Score 241; DB 4; Length 724;
Best Local Similarity 81.2%; Pred. No. 5.9e-62;
Matches 280; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
Qy 10 CTGCTCAGCAGCCTGGGGCTGAACCTGGCAAAACCTGGGGCCTCAGTGAAGATGTCCTGC 69
    |||

```

Db 98 CAGCTGCAGCAGTCTGGAGCTGAGCTGGTGAAGCCTGGGGCTTCAGTGAAGATATCTCTGC 157  
Qy 70 AAGGCTTCTGGCTACACCTTTTACTTAACCTTACTGAGTACTGAGTGAAGTGAAGAGAGGCTGGA 129  
Db 158 AAGGCTTCTGGCTACACCTTTCAAGCTTACTATATACACTGGTGAAGAGAGGCTGGA 217  
Qy 130 CAGGCTCTGAATGATGGATACATTAATCTCTGCACCTGGTTCACCTTCTTACAATCAG 189  
Db 218 CAGGACTTGAAGTGGATGGATGATTTTCTTGGAAATGCTAATCTTAAGTACATGAG 277  
Qy 190 GACTTTTCAGGAGCAGGCGCACTTTGACCGCAGACAAGTCTCTCCACACAGCCTTACATGAG 249  
Db 278 AAGTTCAGGCGCAAGGCCACACTGACTGCAGACAATCTCCAGCAGCCTTACATGAG 337  
Qy 250 CTGACCGCTGACATCTGAGGACTCTTCACTGCTTACTTACTGTCAGAGAGGGGTACGAC 309  
Db 338 ATCAGCAGCCTGAGCTCTGAGGACTCTTTCGGTCTATTTCTGTGCAAGAGAGGGGAGTAT 397  
Qy 310 GGGTTTGACTCTCTGGGGCAAGGACCACTCTCACAGTCTCTCTCA 354  
Db 398 CTCCTTGACTTCTGGGGCAAGGACCACTCTCACAGTCTCTCTCA 442

RESULT 7  
BF582051  
LOCUS 60209230P1 NCI\_CGAP\_Co24 849 bp mRNA linear EST 12-DEC-2000  
DEFINITION mRNA sequence.  
ACCESSION BF582051  
VERSION BF582051.1 GI:11655763  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM9800 row: 1 column: 08  
High quality sequence stop: 732.  
Location/Qualifiers  
1. .849  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4219159"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP Co24"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

FEATURES  
source

Query Match 66.3%; Score 234.6; DB 2; Length 849;  
Best Local Similarity 80.0%; Pred. No. 5.4e-60;  
Matches 276; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
Qy 10 CTGCTCAGCAGCCTGGGCTGAAGTGAAGAGAGGCTTCAAGTGAAGATGTCCTGC 69  
Db 108 CAGCTGCAGCAGTCTGGAGCTGAGCTGGTGAAGCCTGGGACTTCAAGTGAAGATATCTCTGC 167

## ORIGIN

Query Match 65.3%; Score 231; DB 4; Length 729;  
Best Local Similarity 87.8%; Pred. No. 6.5e-59;  
Matches 252; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
Qy 10 CTGCTCAGCAGCCTGGGCTGAAGTGAAGAGAGGCTTCAAGTGAAGATGTCCTGC 69  
Db 99 CAGCTTCAAGGCGCAAGGCCACACTGACTGCAGACAATCTCCAGCAGCCTTACATGAG 158  
Qy 70 AAGGCTTCTGGCTACACCTTTTACTTAACCTTACTGAGTGAAGTGAAGAGAGGCTGGA 129

Qy 70 AAGGCTTCTGGCTACACCTTTTACTTAACCTTACTGAGTGAAGTGAAGAGAGGCTGGA 129  
Db 168 AAGGCTTCTGGCTACACCTTTCAAGCTTACTATATACACTGGTGAAGAGGCTGGA 227  
Qy 130 CAGGCTCTGAATGATGGATACATTAATCTCTGCACCTGGTTCACCTTCTTACAATCAG 189  
Db 228 CAGGACTTGAAGTGGATGGATGATTTATCTTGGAAATATTAATCTTAAGTACATGAG 287  
Qy 190 GACTTTTCAGGAGCAGGCGCACTTTGACCGCAGACAAGTCTCTCCACACAGCCTTACATGAG 249  
Db 288 GTCTTCAAGGCGCAAGGCCACACTGACTGCAGACAATCTCCAGCAGCCTTACATGAG 347  
Qy 250 CTGACCGCTGACATCTGAGGACTCTTCACTGCTTACTTACTGTCAGAGAGGGGTACGAC 309  
Db 348 CTCAAGCCTGAACTCTGAGGACTCTTTCGGTCTATTTCTGTGCAAGAGAGGGGGTAC 407  
Qy 310 GGGTTTGACTCTCTGGGGCAAGGACCACTCTCACAGTCTCTCTCA 354  
Db 408 TATTTTGACTTCTGGGGCAAGGACCACTCTCACAGTCTCTCTCA 452

RESULT 8  
B1100304

LOCUS 602885769P1 NCI\_CGAP\_Kid14 729 bp mRNA linear EST 26-JUN-2001  
DEFINITION 5', mRNA sequence.  
ACCESSION B1100304  
VERSION B1100304.1 GI:14551197  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM1113 row: 1 column: 17  
High quality sequence stop: 727.  
Location/Qualifiers  
1. .729  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5041240"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP Kid14"  
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

FEATURES  
source

Query Match 65.3%; Score 231; DB 4; Length 729;  
Best Local Similarity 87.8%; Pred. No. 6.5e-59;  
Matches 252; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
Qy 10 CTGCTCAGCAGCCTGGGCTGAAGTGAAGAGAGGCTTCAAGTGAAGATGTCCTGC 69  
Db 99 CAGCTTCAAGGCGCAAGGCCACACTGACTGCAGACAATCTCCAGCAGCCTTACATGAG 158  
Qy 70 AAGGCTTCTGGCTACACCTTTTACTTAACCTTACTGAGTGAAGTGAAGAGAGGCTGGA 129

```

Db 159 AAGCTTCTGGCTACACCTTTACTCACTACTGAGATGCACTGGATAAAACAGAGCGCTGGA 218
Qy 130 CAGGCTCTGAATGGATGGATACATTAACTCTGCCACTGGTTCACACTTCTTACAATCAG 189
Db 219 CAGAGTCTGGATGGATGGATACATTAACTCTGACACTGGTCATCTGACTACATCA 278
Qy 190 GACTTTTCAGACAGGGCCACTTTGACCGCAGACAAAGTCTCTCCACACAGCCTCATGCG 249
Db 279 AAATTTCAGGACAAAGCCACATTGACTGCGACACAAATCCTCCAATACAGCCTACATGCA 338
Qy 250 CTGACCGCTGACATCTGAGGACTCTTCACTCTATTACTGTGCAAG 296
Db 339 CTGAGTAGCCTGACATCTGAGGACTCTGCGAGTCTATTACTGTACAAG 385

RESULT 9
BF584024
LOCUS BF584024
DEFINITION 602096264P1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4216550 5',
mRNA sequence.
ACCESSION BF584024
VERSION BF584024.1 GI:11657742
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 698)
AUTHORS NIH-MSC http://msc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9793 row: o column: 15
High quality sequence stop: 682.
FEATURES
    source
        1..698
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clones="IMAGE:4216550"
            /lab_host="DH10B (T1 phage-resistant)"
            /clone_lib="NCI_CGAP Co24"
            /note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.6 kb. Constructed by Life
            Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Query Match 64.5%; Score 228.2; DB 2; Length 698;
Best Local Similarity 78.8%; Pred. No. 4.6e-58;
Matches 272; Conservative 0; Mismatches 73; Indels 0; Gaps . 0;
Qy 10 CTGCTCGACGACCTGGGGCTGAACCTGGCAAACTCTGGGGCTCAGTGAAGATCTCTGCG 69
Db 89 CAGCTGCTCCAGTCTGGACCTGAGCTGGTGAAGCTGGGGCTTCAGTGAAGATATCTCTGC 148
Qy 70 AAGCTTCTGGCTACACCTTTACTAACTACTGGATTCACTGGGTGAACAGAGCGCTTGA 129
Db 149 AAGGCTTTGGGCTACACCTTCAACAGGTACTATATACACTGGGGTGAAGAGCGCTTGA 208
Qy 130 CAGGCTCTGAATGGATGGATACATTAACTCTGCCACTGGTTCACACTTCTTACAATCAG 189
Db 209 CAGGGAATTGTGTGGATTGGATGGATTATCTCTGGAGATGGTATATTAAGTACAATGAG 268

```

```

Qy 190 GACTTTTCAGACAGGGCCACTTTGACCGCAGACAAAGTCTCTCCACACAGCCTCATGCG 249
Db 269 AAGTTCAAGGCAAGACCACTGACTGCGACAAAATCCTCCAGACACAGCCTCATGTTT 328
Qy 250 CTGACCGCTGACATCTGAGGACTCTTCACTCTATTACTGTGCAAGAGGGGTACGAC 309
Db 329 CTCAGCAGCCTGACCTCTGAGGACTCTGCTGTCTATTCTGTGTACAAGAGGGGGGATGG 388
Qy 310 GGGTTTCACTCTCTGGGCGCAAGGACCACTCTCACAGTCTCTCTCA 354
Db 389 GCATTGACTACTGGGCGCAAGGACCACTCTCACAGTCTCTCTCA 433

RESULT 10
CA577954
LOCUS CA577954
DEFINITION K0716F08-5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)
cDNA Library (Long) Mus musculus cDNA clone NIA:K0716F08
IMAGE:30075139 5', mRNA sequence.
ACCESSION CA577954
VERSION CA577954.1 GI:25126345
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 529)
AUTHORS Piao Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
Aiba, K., Taub, D., Longo, D.L., Keller, J. and Ko, M.S.H.,
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)
Unpublished (2001)
JOURNAL Other ESTs: K0716F08-3
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0716 row: F column: 08
Seq primer: M13 Reverse
High quality sequence stop: 529
POLYA=No.
FEATURES
    Location/Qualifiers
        1..529
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6Ncr"
            /db_xref="niaEST:K0716F08-5N"
            /clone="NIA:K0716F08 IMAGE:30075139"
            /tissue_type="Hematopoietic Stem Cell"
            /lin-/c-Kit-/Sca-1-=""
            /dev stage="Age approx. 10 weeks old"
            /lab_host="DH10B"
            /clone_lib="NIA Mouse Hematopoietic Stem Cell"
            /lin-/c-Kit-/Sca-1-=""
            /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
            NotI; Mouse cDNA project by the Laboratory of Genetics,
            National Institute on Aging (NIA), Intramural Research
            Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
            a long-transcript enriched cDNA library (Ref. Genome Res.
            11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were
            obtained from Drs. Dennis Taub, Dan Longo (National
            Institute on Aging, USA), Jonathan Keller (National Cancer
            Institute, USA). Double-stranded cDNAs were synthesized
            with an Oligo(dT) primer [Invitrogen]:
            5'-pGACTAGTCTAGTCGAGCGCGCCCTTTTCTTTTCTTTT-3' from
            0.9 ug of total RNA, treated with T4 DNA polymerase, and
            purified by ethanol-precipitation. The cDNAs were ligated
            to Lona-linker L1-Sal4, purified by phenol/chloroform, and
            separated from free linkers by Centricon 100. Then, the
            cDNAs were amplified by long-range high fidelity PCR using
            Ex Taq polymerase (Takara) with a primer Sal4-S. The

```

products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.1 kb. The library was constructed by Yulan Piao (NIA)."

```
ORIGIN
Query Match      64.1%; Score 227; DB 6; Length 529;
Best Local Similarity 80.9%; Pred. No. 9.8e-58;
Matches 279; Conservative 0; Mismatches 60; Indels 6; Gaps 1;

Qy 10 CTGCTCAGCAGCCTGGGGCTGAACTGCAAAACCTGGGGCTCAGTGAAGATCTCTGC 69
Db 86 CAGCTACAGCAGCTGGACCTGAGCTGGTGAAGCTCTGGGGCTTCAGTGAAGATCTCTGC 145
Qy 70 AAGCTTTCTGGCTACACCTTTTAACTTCACTGGATTCACTGGGTGAAACAGAGGCTTGA 129
Db 146 AAGCTTTCTGGCTACACCTTTTAACTTCACTGGATTCACTGGGTGAAACAGAGGCTTGA 205
Qy 130 CAGGCTCGAAATGGATTGGATACATTAATCTCCCACTGGTTCCTTCAATCAATGAG 189
Db 206 CAGGCACTTGAAGTGGATTGGATGATTTTCTGGAAGTGGTAGTACTTACTCAATGAG 265
Qy 190 GACTTTGAGCAGAGGGCCACTTTGACCGCAGACAGAACTCTCCACACAGGCTTACATGAG 249
Db 266 AAGTTCAAGGGCAAGGGCCACTTACTGTGACAAATCTCTCCACACAGGCTTACATGTTG 325
Qy 250 CTGACCAAGCTGACATCTGAGGACTCTTCACTGTATTACTGTGCAAGAGAGGGGTACGAC 309
Db 326 CTGAGCAGCTGACCTCTGAGGACTCTGCGGTCTATTCTTGTGCAAGAGGGGGTTAC--- 382
Qy 310 GGGTTTGACTCTGGGGCCCAAGGCAACCACTCTCAAGTCTCTCTCA 354
Db 383 ---TTTGACTGCTGGGGCCCAAGGCAACCACTCTCAAGTCTCTCTCA 424
```

```
RESULT 11
BF165486
LOCUS
DEFINITION
60177393F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:4019039 5',
mRNA sequence.
ACCESSION
BF165486
VERSION
BF165486.1 GI:11045851
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 847)
NTH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9270 row: i column: 24
High quality sequence stop: 760.
FEATURES
Location/Qualifiers
1..847
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4019039"
/tissue_type="spontaneous tumor, metastatic to mammary."
```

```
source
1..847
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4019039"
/tissue_type="spontaneous tumor, metastatic to mammary."
```

Stem cell origin."
/lab\_host="DH10B"
/clone\_lib="NCI\_CGAP\_Lu29"
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

```
ORIGIN
Query Match      64.0%; Score 226.6; DB 2; Length 847;
Best Local Similarity 78.6%; Pred. No. 1.5e-57;
Matches 271; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 10 CTGCTCAGCAGCCTGGGGCTGAACTGCAAAACCTGGGGCTCAGTGAAGATCTCTGC 69
Db 77 CAGTTGAAGCAGCTGGAGCTGAGCTGGTGAAGCTCTGGGGCTTCAGTGAAGCTCTCTGC 136
Qy 70 AAGCTTTCTGGCTACACCTTTTAACTTCACTGGATTCACTGGGTGAAACAGAGGCTTGA 129
Db 137 AAGCTTTCTGGCTACACCTTTTAACTTCACTGGATTCACTGGGTGAAACAGAGGCTTGA 196
Qy 130 CAGGCTCGAAATGGATTGGATACATTAATCTCCCACTGGTTCCTTCAATCAATGAG 189
Db 197 CAGGCTTGAAGTGGATGAGTACTTTTCTTCAATGATGATCTTAAATGCAATGAG 256
Qy 190 GACTTTGAGCAGAGGGCCACTTTGACCGCAGACAGAACTCTCCACACAGGCTTACATGAG 249
Db 257 AAGTTCAAGGGCAAGGGCCACTTACTGTGACAAATCTCTCCACACAGGCTTACATGAG 316
Qy 250 CTGACCAAGCTGACATCTGAGGACTCTTCACTGTATTACTGTGCAAGAGAGGGGTACGAC 309
Db 317 CTGAGCAGCTGACCTCTGAGGACTCTGCGGTCTATTCTTGTGCAAGAGGGGGTTAC--- 376
Qy 310 GGGTTTGACTCTGGGGCCCAAGGCAACCACTCTCAAGTCTCTCTCA 354
Db 377 AGCTTGACTACTGGGGCCCAAGGCAACCACTATCAAGTCTCTCTCA 421
```

```
RESULT 12
BF577847
LOCUS
DEFINITION
602092155F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4206540 5',
mRNA sequence.
ACCESSION
BF577847
VERSION
BF577847.1 GI:11651559
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 684)
NTH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9767 row: n column: 13
High quality sequence stop: 672.
FEATURES
Location/Qualifiers
1..684
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4206540"
/lab_host="DH10B (TI phage-resistant)"
```

```
source
1..684
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4206540"
/lab_host="DH10B (TI phage-resistant)"
```

```
/clone_lib="NCI_CGAP_Co24"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match      63.6%; Score 225; DB 2; Length 684;
Best Local Similarity 78.3%; Pred. No. 4.3e-57;
Matches 270; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 10 CTGCTCAGCAGCCTGGGGCTGAACCTGGCAAACTGGGGCTCAGTGAAGATGTCCTGC 69
Db 61 CAGCTGACCAAGTCTGGACCTGAGCTGGTGAAGCTGGGGCTTCAGTGAAGATGTCCTGC 120
Qy 70 AAGGCTTCTGGCTACACCTTTACTTAACCTACTGGATTCACTGGGTGAACAGAGGCTCTGGA 129
Db 121 AAGGCTTCTGGATACACATTCATTAACCTACTTACATTCAGTGGTGAACAGAGGCTCTGGA 180
Qy 130 CAGGGTCTGAATGGATTGGATACATTAATCTGCCACTGGTTCCACTTCTTACAATCAG 189
Db 181 AAGAGCCTTCAGTGGATTGGATATATTTATCTAGCAATGGTGAGACTGGTTACAACAG 240
Qy 190 GACTTTCAGACAGAGGCCACTTTGACCGCAGACAAGTCTCCACCACAGCCTACATCGAG 249
Db 241 AACTTCAAGGCGAAGGCCACATTTGACTGTAGACAAGTCTCCAGCACAGCCTACATGGAC 300
Qy 250 CTGACCAAGCTGACATCTGAGGACTCTTCACTGTATTACTGTGCAAGAGAGGGTACGAC 309
Db 301 CTGGGAGCTGACATCTGAAGATCTGAGCTTATTAATCTTACTGTGCAAGAGACTTTGGTAAC 360
Qy 310 GGTTTGAATCTCTGGGGCCCAAGGACCACTCTCACAGTCTCTCTCA 354
Db 361 TACGGGAGCTACTGGGGTCAAGGAACCTCAGTCACCGTCTCTCTCA 405

RESULT 13
BI647308
LOCUS
DEFINITION
603279436F1 NCI_CGAP_Mam3 920 bp mRNA linear EST 12-SEP-2001
mRNA sequence.
BI647308
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
GI:15561531
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@email.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11810 row: h column: 10
High quality sequence stop: 796.
Location/Qualifiers
1. .920
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5319609"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"

FEATURES
source
1. .501
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6NCr"
/db_xref="niaEST:K0709F01-5N"

/clone_lib="NCI_CGAP_Mam3"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
```





Db	209	GGCCTTGAGTGGATTGGAATGATTTCATCCTTAATAGTGTAGTACTAACTACAATGAGAAG	268
Qy	193	TTTCAGGACAGGGCCACTTTTGACCGCAGACAAAGTCTCTCCACACAGCCTACATGCAGCTG	252
Db	269	TTCAAGAGCAAGGCCACACTGACTGTAGACAAATCCTCCAGCACGGCTACATGCAACTC	328
Qy	253	ACCAGCCTGACATCTGAGGACTCTTCAGTCTATTACTGTGTGCAAGAGAGGGGTACGACGGG	312
Db	329	AGCAGCCTGACATCTGAGGACTCTGCGGTCTATTACTGTGTGCAAGAGGGGGGACAGGGC	388
Qy	313	TTTGACTCCTGGGGCCAAAGGCACCACTCTCACAGTCTCCTCA	354
Db	389	CAGCCTTACTGGGGCCAAAGGACTCTGTGTCACTGTCTCTGCA	430

Search completed: June 4, 2005, 10:23:11  
Job time : 5347.6 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 15:37:35 ; Search time 46.0606 Seconds  
(without alignments)  
41.984 Million cell updates/sec

Title: US-10-089-452-21

Perfect score: 30

Sequence: 1 DTVVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	5	4 AAB86081	Aab86081 H. pylori
2	30	100.0	5	4 AAB86049	Aab86049 H. pylori
3	30	100.0	71	5 ABP64423	Abp64423 Human ORF
4	30	100.0	120	4 AAB86107	Aab86107 H. pylori
5	30	100.0	120	4 AAB86069	Aab86069 H. pylori
6	30	100.0	141	3 AAB03951	Aab03951 Human mes
7	30	100.0	141	8 ADN99711	Adn99711 Novel hum
8	30	100.0	202	5 ABP27888	Abp27888 Streptoco
9	30	100.0	306	4 AAB94314	Aab94314 Human pro
10	30	100.0	306	6 ABU00085	Abu00085 Human nov
11	30	100.0	306	8 ADN99712	Adn99712 Novel hum
12	30	100.0	306	8 ADN14433	Adn14433 Human NF-
13	30	100.0	306	8 ADP25173	Adp25173 PRO polyp
14	30	100.0	316	4 AAM25862	Aam25862 Human pro
15	30	100.0	440	6 ADA54344	Ada54344 Human pro
16	30	100.0	600	6 ABJ25563	Abj25563 Aspergill
17	30	100.0	640	7 ADC31159	Adc31159 Human nov
18	30	100.0	777	6 ABJ26163	Abj26163 Aspergill
19	29	96.7	5	2 AAR85160	Aar85160 Human ONS
20	29	96.7	10	7 ADD35394	Add35394 Human MAB
21	29	96.7	11	6 ABJ26734	Abj26734 VEGF bind
22	29	96.7	11	8. ADK18229	Adk18229 Mouse VEG
23	29	96.7	18	8 ADS00425	Ads00425 Herceptin
24	29	96.7	110	7 ABU62845	Abu62845 Mouse ant
25	29	96.7	116	2 AAR15437	Aar15437 Heavy cha

26	29	96.7	116	6	ABJ26740	Abj26740 VEGF bind
27	29	96.7	116	6	ABO27261	Abo27261 ICAM-1 bi
28	29	96.7	116	6	ABO27269	Abo27269 ICAM-1 bi
29	29	96.7	116	6	ABO27263	Abo27263 ICAM-1 bi
30	29	96.7	116	6	ABO27259	Abo27259 ICAM-1 bi
31	29	96.7	116	6	ABO27255	Abo27255 ICAM-1 bi
32	29	96.7	116	6	ABO27277	Abo27277 Humanised
33	29	96.7	116	6	ABO27273	Abo27273 Murine 1A
34	29	96.7	116	6	ABO27257	Abo27257 ICAM-1 bi
35	29	96.7	116	6	ABO27271	Abo27271 ICAM-1 bi
36	29	96.7	116	6	ABO27267	Abo27267 ICAM-1 bi
37	29	96.7	116	8	ADK18235	Adk18235 Mouse VEG
38	29	96.7	117	7	ADF09936	Adf09936 Antibody
39	29	96.7	120	2	AAR30763	Aar30763 Heavy cha
40	29	96.7	120	2	AAR47040	Aar47040 Sequence
41	29	96.7	120	6	ABR55857	Abr55857 Anti-Her-
42	29	96.7	120	7	ADD35374	Add35374 Human MAB
43	29	96.7	120	8	ADM46698	Adm46698 4D5 heavy
44	29	96.7	120	8	ADO57813	Ado57813 Human HER
45	29	96.7	120	8	ADO57817	Ado57817 Human HER

## ALIGNMENTS

## RESULT 1

AAB86081  
ID AAB86081 standard; peptide; 5 AA.

XX AAB86081;

XX AC

XX 17-JUL-2001 (first entry)

XX DE

XX H. pylori catalase derived antibody HP25/6m/1B5 heavy chain CDR1.

KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;

KW acid-resistant microorganism; complementarity determining region; CDR;

KW feces; heavy chain; light chain.

XX Unidentified.

XX OS

XX WO200127612-A2.

XX PN

XX 19-APR-2001.

XX PD

XX 12-OCT-2000; 2000WO-EP010057.

XX PF

XX 12-OCT-1999; 99EP-00120351.

XX PR 16-MAR-2000; 2000EP-00105592.

XX PR 31-MAR-2000; 2000EP-00107028.

XX PR 10-MAY-2000; 2000EP-00110110.

XX PA

XX (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX PI

XX Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;

XX WPI; 2001-282086/29.

XX DR N-PSDB; AAF88108.

XX DR

XX PT

XX PT

XX PT

XX PT

XX PT

XX PS

XX PS

XX PS

XX PS

XX PS

XX PS

XX PS

XX PS

XX PS

XX PS

XX PS

XX PS

XX PS

XX PS

XX PS

CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence represents a  
 CC complementarity determining region (CDR) from an antibody raised against  
 CC the H. pylori catalase or beta-urease antigen which is used to illustrate  
 CC the method of the invention  
 XX  
 SQ Sequence 5 AA;

Query Match 100.0%; Score 30; DB 4; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||||  
 Db 1 DTYVH 5

RESULT 2  
 AAB86049  
 ID AAB86049 standard; peptide; 5 AA.

XX AAB86049;

XX 17-JUL-2001 (first entry)

XX H. pylori catalase derived antibody HP25/6m/1B5 heavy chain CDR1.

XX Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
 KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
 KW antibacterial; complementarity determining region.

XX Unidentified.

XX WO200127613-A2.

XX 19-APR-2001.

XX 12-OCT-2000; 2000WO-EP010058.

XX 12-OCT-1999; 99EP-00120351.

XX 16-MAR-2000; 2000EP-00105592.

XX 31-MAR-2000; 2000EP-00107028.

XX 10-MAY-2000; 2000EP-00110110.

XX (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX Reiter C, Cullmann G, Heppner P, Ringeis A, Mueller H, Haindl E;

XX WPI; 2001-282087/29.

XX N-PSDB; AAF88051.

PT Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
 PT sample.

XX Claim 19; Page 16; 89pp; German.

XX This invention describes a novel method for detecting, in a mammal,  
 CC infection by an acid-resistant microorganism (A) which comprises reacting  
 CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
 CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
 CC complex is formed with Ag, and the formation of a complex detected. R are  
 CC specific for an Ag which, after passage through the intestines, at least  
 CC in some mammals, retains a native (or corresponding) structure against  
 CC which the mammal produces antibodies (when immunized or infected with  
 CC (A), or its extracts, lysates or derived proteins (or fragments) or  
 CC synthetic peptides). The products of the invention have antibacterial  
 CC activity. The method is used to diagnose infection by Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the  
 CC progress of treatment. Receptors, particularly antibodies, directed

CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an  
 CC infection. This sequence represents a complementarity determining region  
 CC (CDR) from an antibody generated against a Helicobacter pylori antigen  
 CC (catalase or beta-urease) which is used to illustrate the method of the  
 CC invention  
 XX  
 SQ Sequence 5 AA;

Query Match 100.0%; Score 30; DB 4; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||||  
 Db 1 DTYVH 5

RESULT 3  
 ABP64423  
 ID ABP64423 standard; protein; 71 AA.

XX ABP64423;

XX 04-NOV-2002 (first entry)

XX Human ORF793.

XX Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
 KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder.

XX Homo sapiens.

XX US2002082206-A1.

XX 27-JUN-2002.

XX 30-MAY-2001; 2001US-00867550.

XX 30-MAY-2000; 2000US-0208427P.

XX (LEAC/) LEACH M D.

XX (MEHR/) MEHRABAN F.

XX (CONL/) CONLEY P B.

XX (TOPP/) TOPPER J N.

XX (LAWD/) LAW D.

XX Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

XX WPI; 2002-626554/67.

XX N-PSDB; ABQ98986.

XX New polypeptide designated ORFX are present in human atherogenic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including  
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.

XX Claim 10; SEQ ID NO 1586; 78pp; English.

XX The present invention relates to novel human ORFX polypeptides and their  
 CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological  
 CC condition associated with an ORFX-associated disorder, e.g. cancer,

CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/sequence.html?docID=20020082206  
 XX  
 SQ Sequence 71 AA;

Query Match 100.0%; Score 30; DB 5; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTYVH 5  
 |||||  
 DB 20 DTYVH 24

RESULT 4  
 AAB86107  
 ID AAB86107 standard; protein; 120 AA.

XX AAB86107;

DT 17-JUL-2001 (first entry)

DE H. pylori catalase antibody HP25/2m/1BS V region heavy chain.

XX Catalase; beta-urease; antibody; antigen; detection; infection; epitope;  
 KW acid-resistant microorganism; complementarity determining region; CDR;  
 KW feces; heavy chain; light chain.

XX Unidentified.

PN WO200127612-A2.

XX 19-APR-2001.

XX 12-OCT-2000; 2000WO-EP010057.

XX 12-OCT-1999; 99EP-00120351.

PR 16-MAR-2000; 2000EP-00105592.

PR 31-MAR-2000; 2000EP-00107028.

PR 10-MAY-2000; 2000EP-00110110.

XX (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;

XX WPI; 2001-282086/29.

XX N-PSDB; AAF88153.

XX Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing Helicobacter pylori, comprises immunochromatographic detection  
 PT of antigen in feces.

XX Claim 31; Fig 3; 90pp; German.

XX This invention describes a novel method for detecting infection by an  
 CC acid-resistant microorganism (A), in a mammal, using  
 CC immunochromatography. The method is used to diagnose infection by an acid  
 CC resistant microorganism (A), in a mammal, such as Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,  
 CC inexpensive and non-invasive, and may indicate the stage of infection. A  
 CC test strip used in the method may include a filter to eliminate particles  
 CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence represents a  
 CC Helicobacter pylori catalase derived antibody V-region heavy chain  
 CC fragment used to illustrate the method of the invention

XX Sequence 120 AA;

Query Match 100.0%; Score 30; DB 4; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTYVH 5  
 |||||  
 DB 31 DTYVH 35

RESULT 5  
 AAB86069  
 ID AAB86069 standard; protein; 120 AA.

XX AAB86069;

DT 17-JUL-2001 (first entry)

DE H. pylori catalase antibody HP25/6m/1B5 V-region heavy chain.

XX Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
 KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
 KW antibacterial; complementarity determining region; V-region.

XX Unidentified.

PN WO200127613-A2.

XX 19-APR-2001.

XX 12-OCT-2000; 2000WO-EP010058.

XX 12-OCT-1999; 99EP-00120351.

PR 16-MAR-2000; 2000EP-00105592.

PR 31-MAR-2000; 2000EP-00107028.

PR 10-MAY-2000; 2000EP-00110110.

XX (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX Reiter C, Cullmann G, Heppner P, Ringeis A, Mueller H, Haindl E;

XX WPI; 2001-282087/29.

XX N-PSDB; AAF88096.

XX Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
 PT sample.

XX Claim 27; Fig 3; 89pp; German.

XX This invention describes a novel method for detecting, in a mammal,  
 CC infection by an acid-resistant microorganism (A) which comprises reacting  
 CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
 CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
 CC complex is formed with Ag, and the formation of a complex detected. R are  
 CC specific for an Ag which, after passage through the intestines, at least  
 CC in some mammals, retains a native (or corresponding) structure against  
 CC which the mammal produces antibodies (when immunized or infected with  
 CC (A), or its extracts, lysates or derived proteins (or fragments) or  
 CC synthetic peptides). The products of the invention have antibacterial  
 CC activity. The method is used to diagnose infection by Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the  
 CC progress of treatment. Receptors, particularly antibodies, directed  
 CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an  
 CC infection. This sequence represents a Helicobacter pylori anti-catalase  
 CC derived antibody HP25/6m/1B5 V-region heavy chain fragment which is  
 CC described in the method of the invention



```
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
XX WPI; 2004-365511/34.
XX N-PSDB; ADN98927.
XX
XX New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX
XX Claim 14; SEQ ID NO 1311; 532pp; English.
XX
XX The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC protein of the invention.
XX
XX Sequence 141 AA;
SQ
Query Match 100.0%; Score 30; DB 8; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DTYVH 5
Db 50 DTYVH 54
RESULT 8
ABP27888
ID ABP27888 standard; protein; 202 AA.
XX
XX ABP27888;
AC
XX
XX 02-JUL-2002 (first entry)
DT
XX
XX Streptococcus polypeptide SEQ ID NO 4952.
DE
XX
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
XX Streptococcus agalactiae.
OS
XX
XX WO200234771-A2.
FN
XX
XX 02-MAY-2002.
PD
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX
XX WPI; 2004-365511/34.
XX N-PSDB; ADN98927.
XX
XX New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX
XX Claim 14; SEQ ID NO 1311; 532pp; English.
XX
XX The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC protein of the invention.
XX
XX Sequence 141 AA;
SQ
Query Match 100.0%; Score 30; DB 8; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DTYVH 5
Db 50 DTYVH 54
RESULT 8
ABP27888
ID ABP27888 standard; protein; 202 AA.
XX
XX ABP27888;
AC
XX
XX 02-JUL-2002 (first entry)
DT
XX
XX Streptococcus polypeptide SEQ ID NO 4952.
DE
XX
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
XX Streptococcus agalactiae.
OS
XX
XX WO200234771-A2.
FN
XX
XX 02-MAY-2002.
PD
XX 29-OCT-2001; 2001WO-GB004789.
XX
XX 27-OCT-2000; 2000GB-00026333.
XX 24-NOV-2000; 2000GB-00028727.
XX 07-MAR-2001; 2001GB-00005640.
XX (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
XX
XX Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI; 2002-352536/38.
XX N-PSDB; ABN68519.
XX
XX New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
XX Claim 1; Page 3656; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
XX Sequence 202 AA;
SQ
Query Match 100.0%; Score 30; DB 5; Length 202;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DTYVH 5
Db 48 DTYVH 52
RESULT 9
AAB94314
ID AAB94314 standard; protein; 306 AA.
XX
XX AAB94314;
AC
XX
XX 26-JUN-2001 (first entry)
DT
XX
XX Human protein sequence SEQ ID NO:14787.
DE
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
KW Homo sapiens.
XX
XX EP1074617-A2.
FN
XX
XX 07-FEB-2001.
PD
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
XX
XX 27-AUG-1999; 99JP-00300253.
PR
```

```
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX
XX Claim 8; SEQ ID NO 14787; 2537bp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX
XX Sequence 306 AA;
SQ
Query Match 100.0%; Score 30; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DTVVH 5
Db 215 DTVVH 219
|||||
RESULT 10
ABU00085
ID ABU00085 standard; protein; 306 AA.
XX
XX ABU00085;
XX
XX 17-JAN-2003 (first entry)
XX
XX Human novel polypeptide #178.
XX
XX Human; genetic disorder; gene mapping; medical imaging; cancer;
XX neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
XX Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
XX osteoarthritis; periodontal disease; liver fibrosis; viral infection;
XX fungal infection; bacterial infection; autoimmune disease; diabetes;
XX atopic dermatitis.
XX
XX Homo sapiens.
XX
XX WO200274961-A1.
XX
XX 24-OCT-2003; 2003WO-US033947.
XX
PD 26-SEP-2002.
XX
XX 14-MAR-2002; 2002WO-US005109.
XX
XX 15-MAR-2001; 2001US-00810173.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao Q, Ren F;
XX Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
XX Wehrman T, Wang J, Wang D, Drmanac RT;
XX WPI; 2003-040556/03.
XX
XX N-PSDB; ABX05163.
XX
XX New isolated polypeptides and polynucleotides, useful for preventing,
XX treating or ameliorating medical conditions, such as cancer,
XX neurodegenerative disorders, lymphoid cell disorders, bone degenerative
XX disorders, and infections.
XX
XX Claim 9; SEQ ID NO 704; 235pp; English.
XX
XX The invention relates to human polynucleotides and the polypeptides they
XX encode. The polynucleotides and polypeptides are useful in diagnostics,
XX forensics, gene mapping, medical imaging, identification of mutations
XX responsible for genetic disorders or other traits, assessing biodiversity
XX and producing many other types of data and products dependent on DNA and
XX amino acid sequences. They are also useful for preventing, treating or
XX ameliorating medical conditions, such as cancer, neurodegenerative
XX disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
XX disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
XX periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
XX bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
XX Sequences ABG99888-ABG99989 and ABU00010-ABU00433 represent human
XX polypeptides of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied by the European Patent Office
XX
XX Sequence 306 AA;
SQ
Query Match 100.0%; Score 30; DB 6; Length 306;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DTVVH 5
Db 215 DTVVH 219
|||||
RESULT 11
ADN99712
ID ADN99712 standard; protein; 306 AA.
XX
XX ADN99712;
XX
XX 29-JUL-2004 (first entry)
XX
XX Novel human protein sequence #528.
XX
XX anti-inflammatory; dermatological; neuroprotective; immunomodulator;
XX antibacterial; virucide; antipsoriatic; cytostatic; gene therapy;
XX vaccine; inflammatory; CNS; immune disorder; cancer; psoriasis; diabetes;
XX early aging; hormonal imbalance; ischemic heart disease;
XX ulcerative colitis.
XX
XX Homo sapiens.
XX
XX WO2004038003-A2.
XX
XX 06-MAY-2004.
XX
XX 24-OCT-2003; 2003WO-US033947.
XX
```



PR 25-OCT-2002; 2002US-0421061P.  
PR 25-OCT-2002; 2002US-0421080P.  
PR 25-OCT-2002; 2002US-0421552P.  
PR 25-OCT-2002; 2002US-0421614P.  
PR 30-OCT-2002; 2002US-0422177P.  
PR 30-OCT-2002; 2002US-0422178P.  
PR 15-NOV-2002; 2002US-0423555P.  
PR 15-NOV-2002; 2002US-0426384P.  
PR 15-NOV-2002; 2002US-0426394P.  
PR 15-NOV-2002; 2002US-0426430P.  
PR 15-NOV-2002; 2002US-0426916P.  
PR 27-NOV-2002; 2002US-0429224P.  
PR 27-NOV-2002; 2002US-0429275P.  
PR 27-NOV-2002; 2002US-0429302P.  
PR 27-NOV-2002; 2002US-0429326P.  
PR 27-NOV-2002; 2002US-0429651P.  
PR 04-DEC-2002; 2002US-0430645P.  
PR 04-DEC-2002; 2002US-0430651P.  
PR 04-DEC-2002; 2002US-0430657P.  
PR 04-DEC-2002; 2002US-0430663P.  
PR 04-DEC-2002; 2002US-0430668P.  
PR 04-DEC-2002; 2002US-0430684P.  
PR 05-DEC-2002; 2002US-0430937P.  
PR 05-DEC-2002; 2002US-0430965P.  
PR 05-DEC-2002; 2002US-0431458P.  
PR 12-DEC-2002; 2002US-043251P.  
PR 12-DEC-2002; 2002US-0433500P.  
PR 13-DEC-2002; 2002US-0433316P.  
PR 13-DEC-2002; 2002US-0433318P.  
PR 23-DEC-2002; 2002US-0436236P.  
PR 03-JAN-2003; 2003US-0437914P.  
PR 17-JAN-2003; 2003US-0440820P.  
PR 17-JAN-2003; 2003US-0440821P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467192P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476621P.  
PR 09-JUN-2003; 2003US-0476632P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485217P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 08-JUL-2003; 2003US-0485359P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PA Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RP, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;  
PI Wong JGP, Wu G, Zhang H, Zeng C;  
XX WPI; 2004-365511/34.  
DR N-PSDB; ADN98928.  
XX

PT New nucleic acid molecules, useful in preparing a composition for treating or preventing e.g. inflammatory, CNS, bacterial or viral disorders, cancer, psoriasis, diabetes, ischemic heart disease or ulcerative colitis.  
PT Claim 14; SEQ ID NO 1312; 532pp; English.  
XX The invention relates to a nucleic acid molecule comprising a polynucleotide sequence or its complement that encodes a polypeptide. The nucleic acid is useful in preparing a composition for treating or preventing inflammatory, CNS, immune, bacterial or viral disorder, cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic heart disease or ulcerative colitis. This sequence corresponds to a protein of the invention.  
XX Sequence 306 AA;  
SQ

Query Match 100.0%; Score 30; DB 8; Length 306;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DTYVH 5  
Db 215 DTYVH 219  
|||||

RESULT 12  
ADRI4433  
ID ADRI4433 standard; protein; 306 AA.  
XX AC ADRI4433;  
XX DT 21-OCT-2004 (first entry)  
XX DE Human NF-kappaB pathway-associated protein SeqID434.  
XX KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide; antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic; antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic; immunosuppressive; vulnery; gene therapy; immune disorder; inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis; hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour; hyper-IgM syndrome; hypohidrotic ectodermal dysplasia; X-linked anhidrotic ectodermal dysplasia; immunodeficiency; viral infection; HIV-1; Htriv-1; hepatitis B; hepatitis C; EBV; influenza; viral replication; host cell survival; evasion of immune response; rheumatoid arthritis; inflammatory bowel disease; colitis; asthma; atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE; autoimmune disorder; hyper immune activity;  
KW aberrant acute phase response; hypercongenital condition; birth defect; necrotic lesion; wound; organ transplant rejection;  
KW aberrant signal transduction; proliferating disorder; cancer;  
HIV propagation; human.  
OS Homo sapiens.  
XX WO2004065577-A2.  
XX 05-AUG-2004.  
XX 13-JAN-2004; 2004WO-US000798.  
XX 14-JAN-2003; 2003US-0440068P.  
XX 12-MAY-2003; 2003US-0469757P.  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX Nadler SG, Neubauer MG, Feder JN, Carman J;  
XX WPI; 2004-562168/54.  
XX DR N-PSDB; ADRI4432.  
XX New isolated polynucleotides and polypeptides associated with NF-kappaB

pathway, useful for diagnosing, treating, or preventing disorders or diseases associated with NF-kappaB pathway.

Claim 6; SEQ ID NO 434; 237pp; English.

This invention relates to the novel association of protein sequences (and the genes which encode them) to the NF-kappaB pathway. The invention may be useful for the production of compounds with an antiinflammatory, cytostatic, hepatotropic, virucide, antiarthritic, antineumatic, gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic, immunomodulator, cerebroprotective, vasotropic, immunosuppressive or vulnary activity or for gene therapy. The proteins and nucleotides are useful for diagnosing, preventing, treating, or ameliorating conditions or diseases associated with the NF-kappaB pathway. The condition is an immune disorder, an inflammatory disorder, an inflammatory disorder related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis, hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1, hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell survival, evasion of immune responses, rheumatoid arthritis, inflammatory bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper immune activity, disorders related to aberrant acute phase responses, hypercongenital conditions, birth defects, necrotic lesions, wounds, organ transplant rejection, conditions related to organ transplant rejection, disorders related to aberrant signal transduction, proliferating disorders, cancers and HIV propagation in cells infected with other viruses. The present sequence is that of a human protein which is subject to the novel association with the NF-kappaB pathway of the invention. Note: This sequence does not appear in the specification but was obtained by the indexer from Genbank.

Query Match	100.0%	Score 30;	DB 8;	Length 306;
Best Local Similarity	100.0%	Pred. No. 3.2e+02;		
Matches 5;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

RESULT 13	
ADP25173	
ID	ADP25173 standard; protein; 306 AA.
XX	
AC	ADP25173;
XX	
DT	19-NOV-2004 (first entry)
XX	
DE	PRO polypeptide SEQ ID NO:2351.
XX	
KW	PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW	osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW	antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.

DR	WPI; 2004-419628/39.
DR	N-PSDB; ADP25172.
XX	
PT	New PRO polypeptides and polynucleotides, useful for treating e.g.
PT	erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT	renal disease, or demyelinating diseases of the central or peripheral
PT	nervous system.
XX	
PS	Claim 7; SEQ ID NO 2351; 2940pp; English.
XX	
CC	The invention relates to a novel isolated nucleic acid and the PRO
CC	polypeptide encoded by it. A protein of the invention has
CC	antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC	osteoprotic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC	antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC	of the invention may have a use in gene therapy. The PRO polypeptide, its
CC	agonist, antagonist, or antibody that specifically binds to the
CC	polypeptide is useful for treating an immune related disorder such as
CC	systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC	juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC	idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC	vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC	thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC	disease, a demyelinating disease of the central or peripheral nervous
CC	system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC	a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC	disease, infectious or autoimmune chronic active hepatitis, primary
CC	biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC	inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC	disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC	disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC	disease, asthma, allergic rhinitis, atopic dermatitis, food
CC	hypersensitivity, urticaria, an immunologic disease of the lung,
CC	eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC	pneumonitis, a transplantation associated disease, graft rejection or
CC	graft-versus-host disease. The present sequence represents a PRO protein
CC	of the invention.
XX	
SQ	Sequence 306 AA;
	Query Match 100.0%; Score 30; DB 8; Length 306;
	Best Local Similarity 100.0%; Pred. No. 3.2e-02;
	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 DTVVH 5
Db	215 DTVVH 219
RESULT 14	
AA025862	
ID	AA025862 standard; protein; 316 AA.
XX	
AC	AA025862;
XX	
XX	
DT	16-OCT-2001 (first entry)
XX	
DE	Human protein sequence SEQ ID NO:1377.
XX	
KW	Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW	antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW	antibacterial; endocrine; cardiant; central nervous system; virucide;
KW	anti-HIV; fungicide; antimutagen; cardiovascular; antineoplastic; aneemia;
KW	antiregretant; haemostatic; vulnerary; antileuc; osteopathic; aneemia;
KW	dermatological; antiallergic; antiasthmatic; antidiabetic; cycostatic;
KW	neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW	immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW	antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW	cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW	genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW	thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW	allergic rhinitis; diabetes; multiple sclerosis; depression;
KW	Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;

RESULT 14	
AAM25862	
ID	AAM25862 standard; protein; 316 AA.
XX	
AC	AAM25862;
XX	
DT	16-OCT-2001 (first entry)
XX	
DE	Human protein sequence SEQ ID NO:1377.
XX	
KW	Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW	antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW	antibacterial; endocrine; cardiant; central nervous system; virucide;
KW	anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW	antisequestant; haemostatic; vulnery; antiulcer; osteopathic; eczema;
KW	dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW	neuroprotective; antidepressant; neotropic; antiparkinsonian; infection;
KW	immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW	antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW	cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW	genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW	thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW	allergic rhinitis; diabetes; multiple sclerosis; depression;
KW	Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;

neurological disorder.

OS Homo sapiens.

XX EPI293569-A2.

XX WO200153455-A2.

XX PD 26-JUL-2001.

XX PF 22-DEC-2000; 2000WO-US035017.

XX PR 23-DEC-1999; 99US-00471275.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-457603/49.

XX DR N-PSDB; AAH99803.

XX PT Isolated human polynucleotides encoding polypeptides, useful for the

XX PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.

XX PS Claim 20; Page 283; 1217pp; English.

XX CC AAH99166 to AAH99904 encode the human proteins given in AAH25225 to

XX CC AAH25963. The proteins can have activities based on the tissues and cells

XX CC they are expressed in, such as: antiinflammatory; antirheumatic;

XX CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;

XX CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;

XX CC cardiovascular; antianemic; antiaggregant; haemostatic; vulnerary;

XX CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;

XX CC antidabetic; cytostatic; neuroprotective; antidepressant; nootropic;

XX CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides

XX CC encoding them can be used in gene therapy, antisense therapy and vaccine

XX CC production. The proteins and polynucleotides are useful for screening for

XX CC agonists or antagonists of a protein and for the treatment and diagnosis

XX CC of disorders associated with the activity of a protein e.g. inflammation,

XX CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,

XX CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal

XX CC infections, autoimmunity, genetic diseases, haematopoietic disorders,

XX CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,

XX CC osteoporosis, severe combined immunodeficiency, eczema, allergic

XX CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,

XX CC Alzheimer's disease, Parkinson's disease, neurodegenerative and

XX CC neurological disorders

SQ Sequence 316 AA;

Query Match 100.0%; Score 30; DB 4; Length 316;

Best Local Similarity 100.0%; Pred. No. 3.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYPH 5

Db 225 DTYPH 229

RESULT 15

ADA54344

ID ADA54344 standard; protein; 440 AA.

XX AC ADA54344;

XX DT 20-NOV-2003 (first entry)

XX DE Human protein, SEQ ID 1912.

XX KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;

XX KW Gene Therapy; human; secretory protein; membrane proteins; cancer;

XX KW inflammatory disease; osteoporosis; neurological disease.

OS Homo sapiens.

XX EPI293569-A2.

XX PD 19-MAR-2003.

XX PF 21-MAR-2002; 2002EP-00006586.

XX PR 14-SEP-2001; 2001JP-00328381.

XX PR 24-JAN-2002; 2002US-0350435P.

XX PA (HELI-) HELIX RES INST.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Isogai T, Sugiyama T, Otsuki T, Wakanatsu A, Sato H, Ishii S;

XX PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

XX PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX DR WPI; 2003-395539/38.

XX DR N-PSDB; ADA52705.

XX PT New polynucleotides encoding full-length polypeptides, e.g. secretory

XX PT and/or membrane proteins, useful for developing medicines for diseases in

XX PT which the gene is involved, or as target molecules for gene therapy.

XX PS Claim 14; SEQ ID NO 1912; 205pp; English.

XX CC The present invention relates to novel human secretory or membrane

XX CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-

XX CC ADA54071). The coding sequences are useful in the gene therapy of

XX CC diseases caused by abnormalities of the proteins, e.g. cancer,

XX CC inflammatory diseases, osteoporosis or neurological disease.

SQ Sequence 440 AA;

Query Match 100.0%; Score 30; DB 6; Length 440;

Best Local Similarity 100.0%; Pred. No. 4.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYPH 5

Db 257 DTYPH 261

RESULT 16

ABJ25563

ID ABJ25563 standard; protein; 600 AA.

XX AC ABJ25563;

XX DT 16-APR-2003 (first entry)

XX DE Aspergillus fumigatus essential gene protein #221.

XX KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;

XX KW cancer; contamination; biofilm; antibody; immune response.

OS Aspergillus fumigatus.

XX WO200286090-A2.

XX PD 31-OCT-2002.

XX PF 23-APR-2002; 2002WO-US013142.

XX PR 23-APR-2001; 2001US-0285697P.

XX PR 27-APR-2001; 2001US-0287066P.

XX PR 05-JUN-2001; 2001US-0295890P.

XX PR 09-JUL-2001; 2001US-0303899P.

XX PR 31-AUG-2001; 2001US-0316362P.

XX PA (ELIT-) ELITFA PHARM INC.

PI Jiang B, Tishkoff D, Zamudio C, Ershkin AM, Hu W, Lemieux SM;  
DR WPI; 2003-093124/08.  
XX  
XX  
XX New purified or isolated nucleic acids of essential genes of *Aspergillus*  
PT fumigatus, useful for treating or preventing infections by *A. fumigatus*,  
PT or for treating a non-infectious disease in a subject e.g. cancer.  
XX  
XX  
XX Disclosure; Page: 175pp; English.  
PS  
XX The invention relates to novel purified or isolated nucleic acids of  
CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of  
CC organisms are used to treat or prevent infections by a pathogenic  
CC organism such as *A. fumigatus*, to treat a non-infectious disease in a  
CC subject (e.g. cancer), to prevent or contain contamination of an object  
CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a  
CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for  
CC expressing recombinant protein for characterisation, screening or  
CC therapeutic use, as markers for host tissues in which the pathogenic  
CC organisms invade or reside, for comparing with the DNA sequence of *A.*  
CC *fumigatus* to identify duplicated genes or paralogues having the same or  
CC similar biochemical activity and/or function, for comparing with DNA  
CC sequences of other related or distant pathogenic organisms to identify  
CC potential orthologous essential or virulence genes, for selecting and  
CC making oligomers for attachment to a nucleic acid array for examination  
CC of expression patterns, for raising anti-protein antibodies, as an  
CC antigen to raise anti-DNA antibodies or to elicit another immune  
CC response, and for identifying polynucleotides encoding the other protein  
CC with which binding occurs or to identify inhibitors of the binding  
CC interaction. The polypeptides may be used to raise antibodies or to  
CC elicit immune response, as a reagent in assays designed to quantitatively  
CC determine levels of the protein in biological fluids, as a marker for  
CC host tissues in which pathogenic organism invade or reside, and to  
CC isolate correlative receptors or ligands in the case of virulence  
CC factors. This sequence represents a protein of one of the essential genes  
CC of *Aspergillus fumigatus* of the invention  
XX  
XX Sequence 600 AA;  
SQ  
Query Match 100.0%; Score 30; DB 6; Length 600;  
Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DTYVH 5  
Db 414 DTYVH 418  
|||||  
RESULT 17  
ID ADC31159 standard; protein; 640 AA.  
XX  
XX ADC31159;  
XX  
XX 18-DEC-2003 (first entry)  
DT  
XX Human novel polypeptide sequence, SEQ ID NO:1241.  
DE  
XX Human; diagnostic; drug screening; forensics; gene mapping;  
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
KW ulcers; osteoporosis; autoimmune disease; cancer;  
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;  
KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;  
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;  
KW gene therapy; chromosome 8.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2003029271-A2.  
PN  
XX  
XX 10-APR-2003.  
PD  
XX

PF 24-SEP-2002; 2002WO-US030474.  
XX  
XX 24-SEP-2001; 2001US-0324631P.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;  
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;  
PI Haley-Vicente D, Drmanac RT;  
XX  
XX WPI; 2003-371981/35.  
DR N-PSDB; ADC30188.  
DR  
XX New polynucleotide and polypeptide useful for diagnosing, preventing or  
PT treating conditions such as neurodegenerative diseases, anemias, platelet  
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
PT cancer.  
XX  
XX Claim 20; SEQ ID NO 1241; 1185pp; English.  
PS  
XX The invention relates to 971 novel human cDNA sequences (ADC29919-  
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
CC invention also relates to nucleic acid sequences over 99% identical with  
CC the novel human cDNAs. The invention additionally encompasses expression  
CC vectors and host cells comprising a nucleic acid of the invention; the  
CC recombinant production of a polypeptide of the invention; an antibody  
CC against a polypeptide of the invention; a method of detecting  
CC polynucleotides or polypeptides of the invention; and methods of  
CC identifying a compound which binds to a polypeptide of the invention. The  
CC invention further discloses methods of preventing, treating or  
CC ameliorating a medical condition; kits comprising polynucleotide probes  
CC and/or monoclonal antibodies for carrying out the methods of the  
CC invention; methods for the identification of compounds that modulate the  
CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention  
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33394). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are  
CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anaemia, platelet  
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a specifically  
CC claimed human polypeptide sequence of the invention. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 640 AA;  
SQ  
Query Match 100.0%; Score 30; DB 7; Length 640;  
Best Local Similarity 100.0%; Pred. No. 7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DTYVH 5  
Db 457 DTYVH 461  
|||||  
RESULT 18  
ID ABJ26163 standard; protein; 777 AA.  
ABJ26163  
XX  
XX ABJ26163;  
XX  
XX 16-APR-2003 (first entry)  
DT  
XX  
XX *Aspergillus fumigatus* essential gene protein #821.  
DE

XX Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;  
 KW cancer; contamination; biofilm; antibody; immune response.  
 XX Aspergillus fumigatus.  
 XX WO200286090-A2.  
 XX 31-OCT-2002.  
 XX 23-APR-2002; 2002WO-US013142.  
 XX 23-APR-2001; 2001US-0285697P.  
 PR 27-APR-2001; 2001US-0287066P.  
 PR 05-JUN-2001; 2001US-0295890P.  
 PR 09-JUL-2001; 2001US-0303899P.  
 PR 31-AUG-2001; 2001US-0316362P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;  
 XX WPI; 2003-093124/08.  
 XX New purified or isolated nucleic acids of essential genes of Aspergillus  
 PT fumigatus, useful for treating or preventing infections by A. fumigatus,  
 PT or for treating a non-infectious disease in a subject e.g. cancer.  
 XX Disclosure; Page; 175pp; English.  
 XX The invention relates to novel purified or isolated nucleic acids of  
 CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of  
 CC the invention are used to treat or prevent infections by a pathogenic  
 CC organism such as A. fumigatus, to treat a non-infectious disease in a  
 CC subject (e.g. cancer), to prevent or contain contamination of an object  
 CC by A. fumigatus, or to prevent or inhibit formation on a surface of a  
 CC biofilm comprising A. fumigatus. The polynucleotides are useful for  
 CC expressing recombinant protein for characterisation, screening or  
 CC therapeutic use, as markers for host tissues in which the pathogenic  
 CC organisms invade or reside, for comparing with the DNA sequence of A.  
 CC fumigatus to identify duplicated genes or paralogues having the same or  
 CC similar biochemical activity and/or function, for comparing with DNA  
 CC sequences of other related or distant pathogenic organisms to identify  
 CC potential orthologous essential or virulence genes, for selecting and  
 CC making oligomers for attachment to a nucleic acid array for examination  
 CC of expression patterns, for raising anti-protein antibodies, as an  
 CC antigen to raise anti-DNA antibodies or to elicit another immune  
 CC response, and for identifying polynucleotides encoding the other protein  
 CC with which binding occurs or to identify inhibitors of the binding  
 CC interaction. The polypeptides may be used to raise antibodies or to  
 CC elicit immune response, as a reagent in assays designed to quantitatively  
 CC determine levels of the protein in biological fluids, as a marker for  
 CC host tissues in which pathogenic organism invade or reside, and to  
 CC isolate correlative receptors or ligands in the case of virulence  
 CC factors. This sequence represents a protein of one of the essential genes  
 CC of Aspergillus fumigatus of the invention  
 XX Sequence 777 AA;  
 SQ  
 Query Match 100.0%; Score 30; DB 6; Length 777;  
 Best Local Similarity 100.0%; Pred. No. 8.6e+02; Indels 0; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 0;  
 Qy 1 DTYVH 5  
 Db 591 DTYVH 595  
 RESULT 19  
 AAR85160  
 ID AAR85160 standard; protein; 5 AA.  
 XX AAR85160;  
 AC

XX 18-JAN-1996 (first entry)  
 XX Human ONS-M21 antibody heavy variable region CDR 1.  
 DE Human ONS-M21 antibody heavy variable region; chimeric protein;  
 XX complementarity determining region; CDR 1; medulloblastoma; brain tumour;  
 KW treatment; diagnosis.  
 KW Homo sapiens.  
 OS WO9514041-A1.  
 XX 26-MAY-1995.  
 PD 19-OCT-1994; 94WO-JP001763.  
 PF 19-NOV-1993; 93JP-00291078.  
 PR (CHUS ) CHUGAI SEIYAKU KK.  
 XX Ohtomo T, Sato K, Tsuchiya M;  
 XX WPI; 1995-200347/26.  
 DR Reconstituted antibody against human medullo:blastoma cells - contains  
 XX high proportion of human antibody origin and has low antigenicity.  
 PT Claim 11; Page 101; 120pp; Japanese.  
 PS AAR85160-R85162 are human antibody ONS-M21 heavy variable region  
 CC complementarity determining regions (CDRs). They were used in the  
 CC construction of a human/murine chimeric antibody, reactive with human  
 CC medulloblastoma (a brain tumour) cells. The chimeric antibody can be used  
 CC in the diagnosis and treatment of this disease  
 XX Sequence 5 AA;  
 SQ  
 Query Match 96.7%; Score 29; DB 2; Length 5;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 DTYVH 5  
 Db 1 DTYIH 5  
 RESULT 20  
 ADD35394  
 ID ADD35394 standard; peptide; 10 AA.  
 XX ADD35394;  
 AC 15-JAN-2004 (first entry)  
 DT Human MAb4D5-8 VH hypervariable region 1 SEQ ID NO:22.  
 XX antibody heavy chain variable domain; hypervariable region; HER2 binding;  
 KW humanised anti-HER2 antibody; cytostatic; antiinflammatory;  
 KW immunosuppressive; gene therapy; cancer; breast cancer; tumour;  
 KW leukaemia; lymphoid malignancy.  
 XX Homo sapiens.  
 OS WO2003087131-A2.  
 XX 23-OCT-2003.  
 PD 09-APR-2003; 2003WO-US011031.  
 XX 10-APR-2002; 2002US-0371609P.  
 XX (GETH ) GENENTECH INC.  
 PA

XX Lowman HB, Gerstner RB, Carter PJ;  
XX WPI; 2003-845297/78.  
XX  
XX New polypeptide comprising an antibody light or heavy chain variable  
XX variants, specifically anti-HER2 antibody variants, useful for treating  
XX cancer, tumor, inflammatory, immunologic or angiogenic disorders.  
XX  
XX Example 1; SEQ ID NO 22; 96pp; English.  
XX  
XX The present invention describes a polypeptide (I) comprising an antibody  
XX light chain variable domain having hypervariable regions of a sequence of  
XX 109 amino acids (P1, seeADD35373), where one or more amino acids selected  
XX from Q27 (VL), D28 (VL), N30 (VL), T31 (VL), A32 (VL), Y49 (VL), F53 (VL),  
XX Y55 (VL), R66 (VL), H91 (VL), Y92 (VL), and T94 (VL), numbered according to  
XX the Kabat numbering system, are substituted with any amino acid other  
XX than alanine. Also described: (1) an antibody that is capable of binding  
XX to the extracellular domain of HER2, which comprises the hypervariable  
XX regions P1 and/or P2 (see ADD35374); (2) a humanised anti-HER2 antibody  
XX comprising one or more of the substitutions described above; (3) an  
XX article of manufacture comprising a container, a composition contained in  
XX it, and a package insert or label indicating that the composition can be  
XX used to treat cancer characterised by the overexpression of HER2, where  
XX the composition comprises the antibody described above; (4) an antibody  
XX variant of a parent antibody which binds HER2, comprising an amino acid  
XX substitution at position 98 of its heavy chain variable domain, and where  
XX the binding affinity of the antibody variant for HER2 is better than the  
XX binding affinity of the parent antibody for HER2; and (5) isolating high-  
XX affinity variants of humanised anti-HER2 antibody. (I) has cytostatic,  
XX antiinflammatory and immunosuppressive activities, and can be used in  
XX gene therapy. The polypeptide and compositions are useful for treating  
XX cancer, e.g. breast cancer, benign or malignant tumours, leukaemias and  
XX lymphoid malignancies, and other disorders such as neural, glial,  
XX astrocytal, hypothalamic, glandular, macrophagal, epithelial, stromal,  
XX blastocoelec, inflammatory, angiogenic, and immunologic disorders. The  
XX present sequence represent a heavy chain variable domain hypervariable  
XX region of huMAB4D5-8, which is used in the exemplification of the present  
XX invention.  
XX  
XX Sequence 10 AA;

Query Match 96.7%; Score 29; DB 7; Length 10;  
Best Local Similarity 80.0%; Pred. No. 15;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DTVVH 5  
Db 6 DTVIH 10

RESULT 21  
ABJ26734  
ID ABJ26734 standard; peptide; 11 AA.  
XX  
XX AC ABJ26734;

XX  
XX 01-MAY-2003 (first entry)

XX VEGF binding related peptide SEQ ID No 35.

XX Cytostatic; antibody; antigen binding site; VEGF receptor; mitogenesis;  
XX leukaemia cell; vascular endothelial growth factor; tumour;  
XX bispecific antigen-binding protein; mouse; murine.

XX Mus sp.

XX WO2003002144-A1.

XX 09-JAN-2003.

XX 26-JUN-2002; 2002WO-US020332.

PR 26-JUN-2001; 2001US-0301299P.

XX (IMCL-) IMCLONE SYSTEMS INC.

XX Zhu Z;

XX WPI; 2003-201468/19.

XX New bispecific antibodies having antigen-binding sites specific for a  
XX first vascular endothelial growth factor (VEGF) receptor and for a second  
XX VEGF receptor, useful for inhibiting migration of leukemia cells, or for  
XX treating tumors.

XX Claim 19; Page 58; 98pp; English.

XX The invention relates to a novel antibody having a first antigen binding  
XX site specific for a first vascular endothelial growth factor (VEGF)  
XX receptor and a second antigen-binding site specific for a second VEGF  
XX receptor. The bispecific antigen-binding proteins block activation of the  
XX VEGF receptor and are useful for reducing or inhibiting VEGF-induced  
XX cellular functions such as mitogenesis of vascular endothelial cells and  
XX migration of leukaemia cells. The antibodies are useful for treating  
XX tumours and for in vivo or in vitro for investigative and diagnostic  
XX methods. This sequence represents a mouse peptide relating to the  
XX bispecific antibodies that bind to the VEGF receptors of the invention  
XX Sequence 11 AA;

Query Match 96.7%; Score 29; DB 6; Length 11;

Best Local Similarity 80.0%; Pred. No. 17;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DTVVH 5  
Db 7 DTVIH 11

RESULT 22

ADK18229

ID ADK18229 standard; peptide; 11 AA.

XX  
XX AC ADK18229;

XX 06-MAY-2004 (first entry)

XX Mouse VEGF receptor CDR peptide, SEQ ID No 35.

XX antibody; antigen binding site; vascular endothelial growth factor;  
XX VEGF receptor; immunoglobulin; tumour; angiogenesis; cytostatic;  
XX immunostimulant; vaccine; complementarity determining region; CDR; mouse;  
XX murine.

XX Mus sp.

XX WO2004003211-A1.

XX 08-JAN-2004.

XX 24-DEC-2002; 2002WO-US041372.

XX 26-JUN-2002; 2002WO-US020332.

XX (IMCL-) IMCLONE SYSTEMS INC.

XX Zhu Z;

XX WPI; 2004-083065/08.

XX New antibody having a first and second binding sites specific for a first  
XX and second vascular endothelial growth factor (VEGF) receptor, useful for  
XX reducing tumor growth in a mammal and for inhibiting angiogenesis.

XX Claim 19; SEQ ID NO 35; 99pp; English.

XX The invention relates to a novel antibody having a first antigen binding  
 CC site specific for a first vascular endothelial growth factor (VEGF)  
 CC receptor and a second antigen specific for a second VEGF receptor. The  
 CC invention further relates to: an antibody that specifically binds to an  
 CC extracellular domain of a first VEGF receptor and an extracellular domain  
 CC of a second VEGF receptor, where binding of the antibody to the first and  
 CC second VEGF receptor neutralizes activation of that VEGF receptor; a  
 CC method for making the antibody by co-expressing in a host cell a  
 CC recombinant DNA construct encoding a first polypeptide having the first  
 CC immunoglobulin heavy chain domain located to the N-terminus of the second  
 CC immunoglobulin light chain variable domain; a recombinant DNA construct  
 CC encoding a second polypeptide having the second immunoglobulin heavy  
 CC chain variable domain located to the N-terminus of the first  
 CC immunoglobulin light chain variable domain, for time and in a manner  
 CC sufficient to allow expression of the polypeptides and formation of the  
 CC antibody; a method for neutralizing activation of a first VEGF receptor  
 CC and the second VEGF receptor in a cell by treating a cell with the  
 CC antibody cited above; a method for reducing tumour growth in a mammal by  
 CC treating the mammal with the antibody cited above; and a method for  
 CC inhibiting angiogenesis in a mammal by treating the mammal with the  
 CC antibody cited above. The antibody has cytostatic and immunostimulant  
 CC activities. The VEGF receptor antigen can be used to create a vaccine.  
 CC The antibody is useful for reducing tumor growth in a mammal and for  
 CC inhibiting angiogenesis. This sequence represents a VEGF receptor  
 CC complementarity determining region (CDR) peptide of the invention.

XX Sequence 11 AA;

Query Match 96.7%; Score 29; DB 8; Length 11;  
 Best Local Similarity 80.0%; Pred. No. 17;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DTYVH 5  
 Db 7 DTYIH 11  
 |||:|

RESULT 23  
 ADS00425  
 ID ADS00425 standard; peptide; 18 AA.

XX ADS00425;  
 XX  
 DT 16-DEC-2004 (first entry)  
 XX  
 DE Herceptin peptide, SEQ ID 28.  
 XX  
 KW Proteome; human; herceptin; antibody.  
 XX  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 XX Modified-site 1  
 FT /note= "Phosphorylated residue"  
 FT Modified-site 5  
 FT /note= "Phosphorylated residue"

XX WO2004081535-A2.  
 XX  
 XX 23-SEP-2004.  
 XX  
 XX 12-MAR-2004; 2004WO-US007628.

XX 12-MAR-2003; 2003US-0453890P.  
 PR 03-DEC-2003; 2003US-0526832P.  
 XX  
 XX (BION-) BIONOVA CORP.

XX Tsay Y, Wang C;  
 XX  
 XX WPI; 2004-690645/67.

XX

PT Determining a proteome structure profile in an organism, useful for  
 PT assessing the presence, absence or severity of a medical disorder,  
 PT comprises determining a proteome structure profile of proteome(s) in a  
 sample from the organism.

XX Example 5; SEQ ID NO 28; 171pp; English.

XX The present invention relates to a method for determining a proteome  
 CC structure profile in an organism. The method comprises collecting a  
 CC sample from the organism, and determining a proteome structure profile of  
 CC at least one proteome in the sample. The proteome structure profile is  
 CC useful for assessing the presence, absence or severity of a medical  
 CC disorder, for determining the likely outcome resulting to a course of  
 CC treatment, for identifying pharmacological properties of pharmaceuticals,  
 CC or for identifying the etiology and pathophysiology of a medical disorder  
 CC such as disorder associated with infection, with a proliferative disease  
 CC (e.g., cancer), with contact with a toxin (e.g., poison), a condition  
 CC associated with pregnancy, old age, or dementia. The method is also  
 CC useful for identifying the binding partners for proteins or other  
 CC macromolecules, in studies of portal proteins in cells or tissues, as  
 CC well as for understanding many physiological and pathophysiological  
 CC processes and in developing more efficient diagnostic and therapeutic  
 CC means. Furthermore, the method is useful for analysing proteomes from  
 CC particular pathways, organelles, cells, tissues and whole organisms from  
 CC all species. The method is also useful for comparing quaternary  
 CC structural profiles between pathologic and normal specimens to identify  
 CC the quaternary structures that are central to pathogenesis of diseases.  
 CC The present sequence was used to determine the primary and quaternary  
 CC structure information for herceptin antibody, which was used for  
 CC illustrating the method of the invention.

XX Sequence 18 AA;

Query Match 96.7%; Score 29; DB 8; Length 18;  
 Best Local Similarity 80.0%; Pred. No. 28;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DTYVH 5  
 Db 11 DTYIH 15  
 |||:|

RESULT 24  
 ABU62845  
 ID ABU62845 standard; protein; 110 AA.

XX AC ABU62845;  
 XX  
 DT 17-SEP-2003 (first entry)  
 XX  
 DE Mouse anti-LPS monoclonal antibody VH 7-4.

XX WaaP; tyrosine kinase; enzyme; lipopolysaccharide; LPS; HepI; mouse;  
 KW antibody; heavy chain variable region; VH 7-4; monoclonal antibody;  
 XX inner core oligosaccharide; phosphorylation; bacterial infection.

XX Mus sp.  
 OS  
 PN US2003017518-A1.  
 XX  
 XX 23-JAN-2003.  
 XX  
 XX 26-JUN-2002; 2002US-00179851.  
 PR  
 XX 26-JUN-2001; 2001US-0300420P.

XX (LAMJ/) LAM J.  
 PA (ZHAO/) ZHAO X.  
 XX  
 XX Lam J, Zhao X;

XX WPI; 2003-521592/49.  
 DR N-FSDB; ACD2477.

XX Assay for modulators of enzymes involved in phosphorylation of the inner  
PT core oligosaccharide of a lipopolysaccharide, comprises detecting  
PT phosphorylated lipopolysaccharide in a test sample after adding a  
PT specific antibody.  
XX  
XX Disclosure; Fig 14B; 33pp; English.  
XX  
XX The invention relates to assaying for modulators of an enzyme involved in  
CC the phosphorylation of the inner core oligosaccharide (HepI) of  
CC lipopolysaccharides (LPS), comprises: (a) incubating a test sample  
CC comprising the enzyme (e.g. the tyrosine kinase Waap), a candidate  
CC substance and substrates comprising dephosphorylated LPS and a source of  
CC phosphate; (b) adding at least one antibody that binds to phosphorylated  
CC LPS while not binding to dephosphorylated LPS; and (c) detecting  
CC phosphorylated LPS in the test sample by measuring the binding of the  
CC antibody to phosphorylated LPS (by an ELISA assay, enzyme linked  
CC immunosorbent assay). An increase or decrease in the amount of  
CC phosphorylated LPS in the test sample in the presence of the candidate  
CC substance indicates that the candidate substance is a modulator. Also  
CC included is a kit for performing the method. The method is for assaying  
CC for modulators, preferably inhibitors, of an enzyme involved in the  
CC phosphorylation of the inner core oligosaccharide of LPS, wherein a  
CC decrease in the amount of phosphorylated LPS in the test sample in the  
CC presence of the candidate substance indicates that the candidate  
CC substance is an inhibitor. It is useful in conducting a target discovery  
CC business by providing assay systems using the method of the invention for  
CC identifying agents by their ability to modulate and enzyme involved in  
CC the phosphorylation of the inner core oligosaccharide of LPS, conducting  
CC therapeutic profiling of identified agents for efficacy and toxicity in  
CC animals, and licensing the rights to a third party for further drug  
CC development and/or sales or the identified agents or their analogues. The  
CC method does not require the use of radio-labeled substrates. It is  
CC amenable to automation. LPS a major virulence factor for infectious  
CC bacteria such as P. aeruginosa. The present sequence is the heavy chain  
CC variable region from mouse monoclonal antibody 7-4, which recognises the  
CC inner polysaccharide core of P. aeruginosa LPS  
XX  
SQ Sequence 110 AA;  
Query Match 96.7%; Score 29; DB 7; Length 110;  
Best Local Similarity 80.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DTYVH 5  
Db 26 DTYIH 30  
RESULT 25  
AAR15437  
ID AAR15437 standard; protein; 116 AA.  
XX  
XX AAR15437;  
AC  
XX  
DT 25-FEB-1992 (first entry)  
XX  
XX Heavy chain variable region of MAb 1A6.  
DE  
XX HRV; ICAM-1; antigen-binding fragment; inflammation; auto-immune disease.  
KW  
XX Homo sapiens.  
XX  
XX EP459577-A.  
FN  
XX  
PD 04-DEC-1991.  
XX  
XX 25-MAY-1991; 91EP-00201243.  
PF  
XX  
XX 01-JUN-1990; 90US-00532001.  
FR  
XX  
XX (MERI ) MERCK & CO INC.  
PA  
XX

PI Colonna RJ, Condra JH, Tomassini JE, Sardana VV;  
XX  
DR WPI; 1991-355850/49.  
XX  
XX Microbially expressed portions of monoclonal antibody - can block  
PT attachment of rhinovirus ligands to inter-cellular adhesion molecule  
PT (ICAM-1).  
XX  
XX Claim 1; Page 19; 28pp; English.  
PS  
XX This is one of six antibody fragments from MAB's specific for domain 1 of  
CC ICAM-1. MAB 1A6 also specifically blocks the major group of human  
CC rhinovirus from binding to and infecting HeLa cells. The peptide  
CC fragments can be used to treat or prevent rhinovirus infection. See also  
CC AAR15438-R15443  
XX  
SQ Sequence 116 AA;  
Query Match 96.7%; Score 29; DB 2; Length 116;  
Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DTYVH 5  
Db 31 DTYIH 35  
RESULT 26  
ABJ26740  
ID ABJ26740 standard; protein; 116 AA.  
XX  
XX ABJ26740;  
AC  
XX  
DT 01-MAY-2003 (first entry)  
XX  
DE VEGF binding related protein SEQ ID No 41.  
XX  
KW Cytostatic; antibody; antigen binding site; VEGF receptor; mitogenesis;  
KW leukaemia cell; vascular endothelial growth factor; tumour;  
KW bispecific antigen-binding protein; mouse; murine.  
OS Mus sp.  
XX  
XX WO2003002144-A1.  
FN  
XX  
PD 09-JAN-2003.  
XX  
XX 26-JUN-2002; 2002WO-US020332.  
PF  
XX  
XX 26-JUN-2001; 2001US-0301299P.  
FR  
XX  
XX (IMCL-) IMCLONE SYSTEMS INC.  
PA  
XX  
XX Zhu Z;  
PI  
XX  
DR WPI; 2003-201468/19.  
DR  
XX N-PSDB; ABT23320.  
XX  
XX New bispecific antibodies having antigen-binding sites specific for a  
PT first vascular endothelial growth factor (VEGF) receptor and for a second  
PT VEGF receptor, useful for inhibiting migration of leukemia cells, or for  
PT treating tumors.  
XX  
XX Claim 20; Page 59-60; 98pp; English.  
PS  
XX The invention relates to a novel antibody having a first antigen binding  
CC site specific for a first vascular endothelial growth factor (VEGF)  
CC receptor and a second antigen-binding site specific for a second VEGF  
CC receptor. The bispecific antigen-binding proteins block activation of the  
CC VEGF receptor and are useful for reducing or inhibiting VEGF-induced  
CC cellular functions such as mitogenesis of vascular endothelial cells and  
CC migration of leukaemia cells. The antibodies are useful for treating  
CC tumours and for in vivo or in vitro for investigative and diagnostic



CC methods. This sequence represents a mouse protein relating to the  
 CC bispecific antibodies that bind to the VEGF receptors of the invention  
 XX  
 SQ Sequence 116 AA;

Query Match 96.7%; Score 29; DB 6; Length 116;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:|  
 Db 31 DTYIH 35

## RESULT 27

ABO27261  
 ID ABO27261 standard; protein; 116 AA.

XX ABO27261;

AC ABO27261;  
 DT 23-OCT-2003 (revised)  
 DT 11-SEP-2003 (first entry)

XX ICAM-1 binding humanised antibody HumD vH domain.

XX Human; mouse; ICAM-1; intracellular adhesion molecule-1; bacteria;  
 KW human rhinovirus; human coxsackievirus; protozoa; malaria; infection;  
 KW respiratory syncytial virus; fungi; antibody; humanised antibody.

XX Mus sp.  
 OS Homo sapiens.  
 OS Chimeric.

PN US2003035798-A1.

XX 20-FEB-2003.

XX 19-JUL-2001; 2001US-00910483.

XX 30-NOV-1998; 98WO-US025422.

PR 16-AUG-2000; 2000US-00555446.

XX (FANG/) FANG F.

PA (KOH/) KOHLSTAEDT L.

PA (RENO/) RENO J.

XX Fang F, Kohlstaedt L, Reno J;

XX WPI; 2003-503356/47.

XX New humanized antibody binding ICAM-1 and/or inhibits pathogen infection  
 PT of cells expressing ICAM-1, useful for treating infection caused by HRV,  
 PT coxackie A virus, respiratory syncytial virus, bacteria, fungi and  
 PT protozoa.

PS Claim 1; Page 18; 30pp; English.

XX The invention relates to a humanised antibody that binds intracellular  
 CC adhesion molecule-1, ICAM-1 and/or inhibits pathogen infection of cells  
 CC expressing ICAM-1, comprising any of the fully defined Huma to HumI  
 CC sequences. The methods and compositions are useful for inhibiting  
 CC infection and/or progression of infection caused by pathogens such as  
 CC human rhinovirus, coxackie A virus, respiratory syncytial virus,  
 CC bacteria, fungi and protozoa, in particular malaria. The present sequence  
 CC represents the amino acid sequence of ICAM-1 binding humanised antibody  
 CC domain. (Updated on 23-OCT-2003 to standardise OS field)

XX Sequence 116 AA;

Query Match 96.7%; Score 29; DB 6; Length 116;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:|  
 Db 31 DTYIH 35

## RESULT 28

ABO27269  
 ID ABO27269 standard; protein; 116 AA.

XX ABO27269;

XX 23-OCT-2003 (revised)  
 DT 11-SEP-2003 (first entry)

XX ICAM-1 binding humanised antibody HumH vH domain.

XX Human; mouse; ICAM-1; intracellular adhesion molecule-1; bacteria;  
 KW human rhinovirus; human coxsackievirus; protozoa; malaria; infection;  
 KW respiratory syncytial virus; fungi; antibody; humanised antibody.

XX Mus sp.  
 OS Homo sapiens.  
 OS Chimeric.

PN US2003035798-A1.

XX 20-FEB-2003.

XX 19-JUL-2001; 2001US-00910483.

XX 30-NOV-1998; 98WO-US025422.

PR 16-AUG-2000; 2000US-00555446.

XX (FANG/) FANG F.

PA (KOH/) KOHLSTAEDT L.

PA (RENO/) RENO J.

XX Fang F, Kohlstaedt L, Reno J;

XX WPI; 2003-503356/47.

XX New humanized antibody binding ICAM-1 and/or inhibits pathogen infection  
 PT of cells expressing ICAM-1, useful for treating infection caused by HRV,  
 PT coxackie A virus, respiratory syncytial virus, bacteria, fungi and  
 PT protozoa.

PS Claim 1; Page 19; 30pp; English.

XX The invention relates to a humanised antibody that binds intracellular  
 CC adhesion molecule-1, ICAM-1 and/or inhibits pathogen infection of cells  
 CC expressing ICAM-1, comprising any of the fully defined Huma to HumI  
 CC sequences. The methods and compositions are useful for inhibiting  
 CC infection and/or progression of infection caused by pathogens such as  
 CC human rhinovirus, coxackie A virus, respiratory syncytial virus,  
 CC bacteria, fungi and protozoa, in particular malaria. The present sequence  
 CC represents the amino acid sequence of ICAM-1 binding humanised antibody  
 CC domain. (Updated on 23-OCT-2003 to standardise OS field)

XX Sequence 116 AA;

Query Match 96.7%; Score 29; DB 6; Length 116;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:|  
 Db 31 DTYIH 35

## RESULT 29

ABO27263  
 ID ABO27263 standard; protein; 116 AA.

XX

AC ABO27263;  
XX  
XX 23-OCT-2003 (revised)  
DT 11-SEP-2003 (first entry)  
XX  
XX  
XX ICAM-1 binding humanised antibody HumE vH domain.  
XX  
XX Human; mouse; ICAM-1; intracellular adhesion molecule-1; bacteria;  
KW human rhinovirus; human coxsackievirus; protozoa; malaria; infection;  
KW respiratory syncytial virus; fungi; antibody; humanised antibody.  
XX  
XX Mus sp.  
OS Homo sapiens.  
OS Chimeric.  
XX  
XX US2003035798-A1.  
XX  
XX 20-FEB-2003.  
XX  
XX 19-JUL-2001; 2001US-00910483.  
XX  
XX 30-NOV-1998; 98WO-US025422.  
XX  
XX 16-AUG-2000; 2000US-00555446.  
XX  
XX (FANG/) FANG F.  
PA (KOHL/) KOHLSTAEDET L.  
PA (RENO/) RENO J.  
XX  
XX Fang F, Kohlstaedt L, Reno J;  
PI WPI; 2003-503356/47.  
XX  
XX New humanized antibody binding ICAM-1 and/or inhibits pathogen infection  
DR of cells expressing ICAM-1, useful for treating infection caused by HRV,  
XX coxackie A virus, respiratory syncytial virus, bacteria, fungi and  
XX protozoa.  
XX  
XX Claim 1; Page 18; 30pp; English.  
XX  
XX The invention relates to a humanised antibody that binds intracellular  
CC adhesion molecule-1, ICAM-1 and/or inhibits pathogen infection of cells  
CC expressing ICAM-1, comprising any of the fully defined Huma to Humi  
CC sequences. The methods and compositions are useful for inhibiting  
CC infection and/or progression of infection caused by pathogens such as  
CC human rhinovirus, coxackie A virus, respiratory syncytial virus,  
CC bacteria, fungi and protozoa, in particular malaria. The present sequence  
CC represents the amino acid sequence of ICAM-1 binding humanised antibody  
CC domain. (Updated on 23-OCT-2003 to standardise OS field)  
XX  
XX Sequence 116 AA;  
SQ  
Query Match 96.7%; Score 29; DB 6; Length 116;  
Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DTIVH 5  
Db 31 DTIVH 35  
RESULT 30  
ABO27259  
ID ABO27259 standard; protein; 116 AA.  
XX  
XX ABO27259;  
XX  
XX 23-OCT-2003 (revised)  
DT 11-SEP-2003 (first entry)  
XX  
XX ICAM-1 binding humanised antibody HumC vH domain.  
XX  
XX Human; mouse; ICAM-1; intracellular adhesion molecule-1; bacteria;  
KW human rhinovirus; human coxsackievirus; protozoa; malaria; infection;

KW respiratory syncytial virus; fungi; antibody; humanised antibody.  
XX  
XX Mus sp.  
OS Homo sapiens.  
OS Chimeric.  
XX  
XX US2003035798-A1.  
XX  
XX 20-FEB-2003.  
XX  
XX 19-JUL-2001; 2001US-00910483.  
XX  
XX 30-NOV-1998; 98WO-US025422.  
XX  
XX 16-AUG-2000; 2000US-00555446.  
XX  
XX (FANG/) FANG F.  
PA (KOHL/) KOHLSTAEDET L.  
PA (RENO/) RENO J.  
XX  
XX Fang F, Kohlstaedt L, Reno J;  
PI WPI; 2003-503356/47.  
XX  
XX New humanized antibody binding ICAM-1 and/or inhibits pathogen infection  
DR of cells expressing ICAM-1, useful for treating infection caused by HRV,  
XX coxackie A virus, respiratory syncytial virus, bacteria, fungi and  
XX protozoa.  
XX  
XX Claim 1; Page 18; 30pp; English.  
XX  
XX The invention relates to a humanised antibody that binds intracellular  
CC adhesion molecule-1, ICAM-1 and/or inhibits pathogen infection of cells  
CC expressing ICAM-1, comprising any of the fully defined Huma to Humi  
CC sequences. The methods and compositions are useful for inhibiting  
CC infection and/or progression of infection caused by pathogens such as  
CC human rhinovirus, coxackie A virus, respiratory syncytial virus,  
CC bacteria, fungi and protozoa, in particular malaria. The present sequence  
CC represents the amino acid sequence of ICAM-1 binding humanised antibody  
CC domain. (Updated on 23-OCT-2003 to standardise OS field)  
XX  
XX Sequence 116 AA;  
SQ  
Query Match 96.7%; Score 29; DB 6; Length 116;  
Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DTIVH 5  
Db 31 DTIVH 35  
RESULT 31  
ABO27255  
ID ABO27255 standard; protein; 116 AA.  
XX  
XX ABO27255;  
XX  
XX 23-OCT-2003 (revised)  
DT 11-SEP-2003 (first entry)  
XX  
XX ICAM-1 binding humanised antibody Huma vH domain.  
XX  
XX Human; mouse; ICAM-1; intracellular adhesion molecule-1; bacteria;  
KW human rhinovirus; human coxsackievirus; protozoa; malaria; infection;  
KW respiratory syncytial virus; fungi; antibody; humanised antibody.  
XX  
XX Mus sp.  
OS Homo sapiens.  
OS Chimeric.  
XX  
XX US2003035798-A1.  
XX  
XX 20-FEB-2003.

XX PF 19-JUL-2001; 2001US-00910483.  
 XX PR 30-NOV-1998; 98WO-US025422.  
 XX PR 16-AUG-2000; 2000US-00555446.  
 XX PA (FANG/) FANG F.  
 XX PA (KOH/) KOHLSTAEDE L.  
 XX PA (RENO/) RENO J.  
 XX PI Fang F, Kohlstaedt L, Reno J;  
 XX DR WPI; 2003-503356/47.  
 XX PT New humanized antibody binding ICAM-1 and/or inhibits pathogen infection  
 XX PT of cells expressing ICAM-1, useful for treating infection caused by HRV,  
 XX PT coxackie A virus, respiratory syncytial virus, bacteria, fungi and  
 XX PT protozoa.  
 XX PS Claim 1; Page 17; 30pp; English.  
 XX CC The invention relates to a humanised antibody that binds intracellular  
 XX CC adhesion molecule-1, ICAM-1 and/or inhibits pathogen infection of cells  
 XX CC expressing ICAM-1, comprising any of the fully defined Huma to Humi  
 XX CC sequences. The methods and compositions are useful for inhibiting  
 XX CC infection and/or progression of infection caused by pathogens such as,  
 XX CC human rhinovirus, coxackie A virus, respiratory syncytial virus,  
 XX CC bacteria, fungi and protozoa, in particular malaria. The present sequence  
 XX CC represents the amino acid sequence of ICAM-1 binding humanised antibody  
 XX CC domain. (Updated on 23-OCT-2003 to standardise OS field)  
 XX SQ Sequence 116 AA;

Query Match 96.7%; Score 29; DB 6; Length 116;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTYVH 5  
 DB 31 DTYIH 35  
 |||:|

RESULT 32  
 ABO27277  
 ID ABO27277 standard; protein; 116 AA.  
 XX AC ABO27277;  
 XX AC  
 XX DT 23-OCT-2003 (revised)  
 XX DT 11-SEP-2003 (first entry)  
 XX DE Humanised 1A6 (Hum19) antibody vH domain.  
 XX DE  
 XX KW Human; mouse; 1A6; infection; human rhinovirus; human coxsackievirus;  
 XX KW protozoa; malaria; respiratory syncytial virus; bacteria; fungi;  
 XX KW humanised antibody; antibody.  
 XX OS Mus sp.  
 XX OS Homo sapiens.  
 XX OS Chimeric.  
 XX PN US2003035798-A1.  
 XX PD 20-FEB-2003.  
 XX PF 19-JUL-2001; 2001US-00910483.  
 XX PR 30-NOV-1998; 98WO-US025422.  
 XX PR 16-AUG-2000; 2000US-00555446.  
 XX PA (FANG/) FANG F.  
 XX PA (KOH/) KOHLSTAEDE L.  
 XX PA (RENO/) RENO J.

XX PI Fang F, Kohlstaedt L, Reno J;  
 XX DR WPI; 2003-503356/47.  
 XX PT New humanized antibody binding ICAM-1 and/or inhibits pathogen infection  
 XX PT of cells expressing ICAM-1, useful for treating infection caused by HRV,  
 XX PT coxackie A virus, respiratory syncytial virus, bacteria, fungi and  
 XX PT protozoa.  
 XX PS Example 2; Fig 3; 30pp; English.  
 XX CC The invention relates to a humanised antibody that binds intracellular  
 XX CC adhesion molecule-1, ICAM-1 and/or inhibits pathogen infection of cells  
 XX CC expressing ICAM-1, comprising any of the fully defined Huma to Humi  
 XX CC sequences. The methods and compositions are useful for inhibiting  
 XX CC infection and/or progression of infection caused by pathogens such as  
 XX CC human rhinovirus, coxackie A virus, respiratory syncytial virus,  
 XX CC bacteria, fungi and protozoa, in particular malaria. The present sequence  
 XX CC represents the amino acid sequence of a humanised antibody 1A6 domain.  
 XX CC (Updated on 23-OCT-2003 to standardise OS field)  
 XX SQ Sequence 116 AA;

Query Match 96.7%; Score 29; DB 6; Length 116;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTYVH 5  
 DB 31 DTYIH 35  
 |||:|

RESULT 33  
 ABO27273  
 ID ABO27273 standard; protein; 116 AA.  
 XX AC ABO27273;  
 XX DT 11-SEP-2003 (first entry)  
 XX DE Murine 1A6 antibody vH domain.  
 XX KW Mouse; 1A6; infection; human rhinovirus; human coxsackievirus; fungi;  
 XX KW protozoa; malaria; respiratory syncytial virus; bacteria; antibody.  
 XX OS Mus sp.  
 XX PN US2003035798-A1.  
 XX PD 20-FEB-2003.  
 XX PF 19-JUL-2001; 2001US-00910483.  
 XX PR 30-NOV-1998; 98WO-US025422.  
 XX PR 16-AUG-2000; 2000US-00555446.  
 XX PA (FANG/) FANG F.  
 XX PA (KOH/) KOHLSTAEDE L.  
 XX PA (RENO/) RENO J.  
 XX PI Fang F, Kohlstaedt L, Reno J;  
 XX DR WPI; 2003-503356/47.  
 XX PT New humanized antibody binding ICAM-1 and/or inhibits pathogen infection  
 XX PT of cells expressing ICAM-1, useful for treating infection caused by HRV,  
 XX PT coxackie A virus, respiratory syncytial virus, bacteria, fungi and  
 XX PT protozoa.  
 XX PS Example 1; Fig 1; 30pp; English.  
 XX CC The invention relates to a humanised antibody that binds intracellular

CC adhesion molecule-1, ICAM-1 and/or inhibits pathogen infection of cells  
 CC expressing ICAM-1, comprising any of the fully defined Huma to Humi  
 CC sequences. The methods and compositions are useful for inhibiting  
 CC infection and/or progression of infection caused by pathogens such as  
 CC human rhinovirus, coxackie A virus, respiratory syncytial virus,  
 CC bacteria, fungi and protozoa, in particular malaria. The present sequence  
 CC represents the amino acid sequence of a mouse antibody 1A6 domain  
 XX  
 SQ Sequence 116 AA;

Query Match 96.7%; Score 29; DB 6; Length 116;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYYH 5  
 |||:|  
 Db 31 DTYYH 35

RESULT 34  
 ABO27257  
 ID ABO27257 standard; protein; 116 AA.

XX AC ABO27257;

DT 23-OCT-2003 (revised)  
 DT 11-SEP-2003 (first entry)

XX ICAM-1 binding humanised antibody HumB vH domain.

XX Human; mouse; ICAM-1; intracellular adhesion molecule-1; bacteria;  
 KW human rhinovirus; human coxsackievirus; protozoa; malaria; infection;  
 KW respiratory syncytial virus; fungi; antibody; humanised antibody.

XX Mus sp.  
 OS Homo sapiens.  
 OS Chimeric.

XX US2003035798-A1.

XX 20-FEB-2003.

XX 19-JUL-2001; 2001US-00910483.

XX 30-NOV-1998; 98WO-US025422.

XX 16-AUG-2000; 2000US-00555446.

XX (FANG/) FANG F.  
 PA (KOHL/) KOHLSTAEDT L.  
 PA (RENO/) RENO J.

XX Fang F, Kohlstaedt L, Reno J;

XX WPI; 2003-503356/47.

XX New humanized antibody binding ICAM-1 and/or inhibits pathogen infection  
 PT of cells expressing ICAM-1, useful for treating infection caused by HRV,  
 PT coxackie A virus, respiratory syncytial virus, bacteria, fungi and  
 PT protozoa.

XX Claim 1; Page 17; 30pp; English.

XX The invention relates to a humanised antibody that binds intracellular  
 CC adhesion molecule-1, ICAM-1 and/or inhibits pathogen infection of cells  
 CC expressing ICAM-1, comprising any of the fully defined Huma to Humi  
 CC sequences. The methods and compositions are useful for inhibiting  
 CC infection and/or progression of infection caused by pathogens such as  
 CC human rhinovirus, coxackie A virus, respiratory syncytial virus,  
 CC bacteria, fungi and protozoa, in particular malaria. The present sequence  
 CC represents the amino acid sequence of ICAM-1 binding humanised antibody  
 CC domain. (Updated on 23-OCT-2003 to standardise OS field)

XX Sequence 116 AA;

Query Match 96.7%; Score 29; DB 6; Length 116;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYYH 5  
 |||:|  
 Db 31 DTYYH 35

RESULT 35

ABO27271  
 ID ABO27271 standard; protein; 116 AA.

XX AC ABO27271;

XX 23-OCT-2003 (revised)  
 DT 11-SEP-2003 (first entry)

XX ICAM-1 binding humanised antibody HumI vH domain.

XX Human; mouse; ICAM-1; intracellular adhesion molecule-1; bacteria;  
 KW human rhinovirus; human coxsackievirus; protozoa; malaria; infection;  
 KW respiratory syncytial virus; fungi; antibody; humanised antibody.

XX Mus sp.  
 OS Homo sapiens.  
 OS Chimeric.

XX US2003035798-A1.

XX 20-FEB-2003.

XX 19-JUL-2001; 2001US-00910483.

XX 30-NOV-1998; 98WO-US025422.

XX 16-AUG-2000; 2000US-00555446.

XX (FANG/) FANG F.  
 PA (KOHL/) KOHLSTAEDT L.  
 PA (RENO/) RENO J.

XX Fang F, Kohlstaedt L, Reno J;

XX WPI; 2003-503356/47.

XX New humanized antibody binding ICAM-1 and/or inhibits pathogen infection  
 PT of cells expressing ICAM-1, useful for treating infection caused by HRV,  
 PT coxackie A virus, respiratory syncytial virus, bacteria, fungi and  
 PT protozoa.

XX Claim 1; Page 20; 30pp; English.

XX The invention relates to a humanised antibody that binds intracellular  
 CC adhesion molecule-1, ICAM-1 and/or inhibits pathogen infection of cells  
 CC expressing ICAM-1, comprising any of the fully defined Huma to Humi  
 CC sequences. The methods and compositions are useful for inhibiting  
 CC infection and/or progression of infection caused by pathogens such as  
 CC human rhinovirus, coxackie A virus, respiratory syncytial virus,  
 CC bacteria, fungi and protozoa, in particular malaria. The present sequence  
 CC represents the amino acid sequence of ICAM-1 binding humanised antibody  
 CC domain. (Updated on 23-OCT-2003 to standardise OS field)

XX Sequence 116 AA;

Query Match 96.7%; Score 29; DB 6; Length 116;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYYH 5  
 |||:|  
 Db 31 DTYYH 35

RESULT 36  
 ABO27267  
 ID ABO27267 standard; protein; 116 AA.  
 XX AC ABO27267;  
 XX DT 23-OCT-2003 (revised)  
 XX DT 11-SEP-2003 (first entry)  
 XX DE ICAM-1 binding humanised antibody HumG vH domain.  
 XX KW Human; mouse; ICAM-1; intracellular adhesion molecule-1; bacteria;  
 KW human rhinovirus; human coxsackievirus; protozoa; malaria; infection;  
 KW respiratory syncytial virus; fungi; antibody; humanised antibody.  
 XX OS Mus sp.  
 OS Homo sapiens.  
 OS Chimeric.  
 XX US2003035798-A1.  
 XX PD 20-FEB-2003.  
 XX PF 19-JUL-2001; 2001US-00910483.  
 XX PR 30-NOV-1998; 98WO-US025422.  
 PR 16-AUG-2000; 2000US-00555446.  
 XX (FANG//) FANG F.  
 PA (KOH//) KOHLSTAEDT L.  
 PA (RENO//) RENO J.  
 XX PI Fang F, Kohlstaedt L, Reno J;  
 XX WPI; 2003-503356/47.  
 XX New humanized antibody binding ICAM-1 and/or inhibits pathogen infection  
 PT of cells expressing ICAM-1, useful for treating infection caused by HRV,  
 PT coxackie A virus, respiratory syncytial virus, bacteria, fungi and  
 PT protozoa.  
 XX Claim 1; Page 19; 30pp; English.  
 XX The invention relates to a humanised antibody that binds intracellular  
 CC adhesion molecule-1, ICAM-1 and/or inhibits pathogen infection of cells  
 CC expressing ICAM-1, comprising any of the fully defined HumA to HumI  
 CC sequences. The methods and compositions are useful for inhibiting  
 CC infection and/or progression of infection caused by pathogens such as  
 CC human rhinovirus, coxackie A virus, respiratory syncytial virus,  
 CC bacteria, fungi and protozoa, in particular malaria. The present sequence  
 CC represents the amino acid sequence of ICAM-1 binding humanised antibody  
 CC domain. (Updated on 23-OCT-2003 to standardise OS field)  
 XX Sequence 116 AA;  
 SQ

Query Match 96.7%; Score 29; DB 6; Length 116;  
 Best Local Similarity 80.0%; Pred. NO. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DTYVH 5  
 |||:  
 Db 31 DTYIH 35

RESULT 37  
 ADK18235  
 ID ADK18235 standard; protein; 116 AA.  
 XX AC ADK18235;  
 XX DT 06-MAY-2004 (first entry)  
 XX

DE Mouse VEGF receptor heavy chain variable domain peptide, SEQ ID NO 41.  
 XX antibody; antigen binding site; vascular endothelial growth factor;  
 KW VEGF receptor; immunoglobulin; tumour; angiogenesis; cytostatic;  
 KW immunostimulant; vaccine; complementarity determining region; variable;  
 KW heavy; mouse; murine.  
 XX Mus.sp  
 OS WO2004003211-A1.  
 XX PD 08-JAN-2004.  
 XX PF 24-DEC-2002; 2002WO-US041372.  
 XX PR 26-JUN-2002; 2002WO-US020332.  
 PR (IMCL-) IMCLONE SYSTEMS INC.  
 PA PA  
 PA Zhu Z;  
 PI WPI; 2004-083065/08.  
 XX New antibody having a first and second binding sites specific for a first  
 PT and second vascular endothelial growth factor (VEGF) receptor, useful for  
 PT reducing tumor growth in a mammal and for inhibiting angiogenesis.  
 XX Claim 20; SEQ ID NO 41; 99pp; English.  
 XX The invention relates to a novel antibody having a first antigen binding  
 CC site specific for a first vascular endothelial growth factor (VEGF)  
 CC receptor and a second antigen specific for a second VEGF receptor. The  
 CC invention further relates to: an antibody that specifically binds to an  
 CC extracellular domain of a first VEGF receptor and an extracellular domain  
 CC of a second VEGF receptor, where binding of the antibody to the first and  
 CC second VEGF receptor neutralizes activation of that VEGF receptor; a  
 CC method for making the antibody by co-expressing in a host cell a  
 CC recombinant DNA construct encoding a first polypeptide having the first  
 CC immunoglobulin heavy chain domain located to the N-terminus of the second  
 CC immunoglobulin light chain variable domain; a recombinant DNA construct  
 CC encoding a second polypeptide having the second immunoglobulin heavy  
 CC chain variable domain located to the N-terminus of the first  
 CC immunoglobulin light chain variable domain, for time and in a manner  
 CC sufficient to allow expression of the polypeptides and formation of the  
 CC antibody; a method for neutralizing activation of a first VEGF receptor  
 CC and the second VEGF receptor in a cell by treating a cell with the  
 CC antibody cited above; a method for reducing tumour growth in a mammal by  
 CC treating the mammal with the antibody cited above; and a method for  
 CC inhibiting angiogenesis in a mammal by treating the mammal with the  
 CC antibody cited above. The antibody has cytostatic and immunostimulant  
 CC activities. The VEGF receptor antigen can be used to create a vaccine.  
 CC The antibody is useful for reducing tumor growth in a mammal and for  
 CC inhibiting angiogenesis. This sequence represents a VEGF receptor  
 CC variable heavy chain peptide of the invention.  
 XX Sequence 116 AA;  
 SQ

Query Match 96.7%; Score 29; DB 8; Length 116;  
 Best Local Similarity 80.0%; Pred. NO. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DTYVH 5  
 |||:  
 Db 31 DTYIH 35

RESULT 38  
 ADF09936  
 ID ADF09936 standard; protein; 117 AA.  
 XX AC ADF09936;  
 XX DT 12-FEB-2004 (first entry)  
 XX

XX Antibody heavy chain variable region 1FVC(3-66) .  
 XX Antibody; stability; solubility; antigen binding affinity;  
 KW variable region; human.  
 XX Homo sapiens.  
 XX WO2003074679-A2.  
 XX 12-SEP-2003.  
 XX 03-MAR-2003; 2003WO-US006598.  
 PF 01-MAR-2002; 2002US-0360843P.  
 PR 29-MAY-2002; 2002US-0384197P.  
 XX (XENC-) XENCOR.  
 XX Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;  
 XX WPI; 2003-722066/68.  
 XX Computer optimization of physicochemical properties of antibodies  
 PT comprises analyzing the interactions of amino acids at variable  
 PT positions.  
 XX Disclosure; Fig 2a; 135pp; English.  
 XX The present invention relates to a method for optimizing at least one  
 CC physico-chemical property of an antibody by a computational screening  
 CC method. The method comprises: receiving a template antibody structure;  
 CC selecting at least one variable position belonging to the antibody  
 CC structure; selecting at least one amino acid to be considered at the  
 CC variable position(s); analyzing the interaction of each selected amino  
 CC acid at each variable position with at least part of the remainder of the  
 CC antibody, including the selected amino acids at other variable positions;  
 CC and identifying a set of at least one antibody sequence with at least one  
 CC optimized physico-chemical property. The method is useful for optimizing  
 CC the physico-chemical properties of an antibody, especially the stability,  
 CC solubility, or antigen binding affinity. The optimized antibody may be  
 CC useful for treating a patient. The present sequence is an antibody  
 CC variable region sequence used to illustrate the invention.  
 XX Sequence 117 AA;  
 SQ

Query Match 96.7%; Score 29; DB 7; Length 117;  
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DTYVH 5  
 |||:  
 Db 31 DTYIH 35

RESULT 39  
 AAR30763  
 ID AAR30763 standard; protein; 120 AA.  
 XX AAR30763;  
 AC AAR30763;

DT 25-MAR-2003 (revised)  
 DT 12-MAY-1993 (first entry)

XX Heavy chain variable domain of humanised muMAB4D5.  
 XX Humanisation; rapid; monoclonal antibody; murine.

OS Mus musculus.  
 XX

PN WO9222653-A1.  
 XX

FD 23-DEC-1992.

XX 15-JUN-1992; 92WO-US005126.  
 XX 14-JUN-1991; 91US-00715272.  
 PR (GETH ) GENENTECH INC.  
 PA Carter PJ, Presta LG;  
 PI WPI; 1993-018139/02.  
 DR Humanisation of antibodies - by molecular modelling of the variable  
 XX domains and alteration by gene conversion mutagenesis.  
 PT Claim 14; Page 8; 126pp; English.  
 PS The sequence is that of the heavy chain variable domain of a humanised  
 CC version of murine monoclonal antibody muMAB4D5. It was produced using a  
 CC method that allows rapid humanisation of antibodies (Ab's) and requires  
 CC less DNA synthesis than previous methods. It has the desired binding and  
 CC other characteristics and activities but is less antigenic in humans than  
 CC non-human Ab's. (Updated on 25-MAR-2003 to correct PN field.)  
 XX Sequence 120 AA;  
 SQ

Query Match 96.7%; Score 29; DB 2; Length 120;  
 Best Local Similarity 80.0%; Pred. No. 2e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DTYVH 5  
 |||:  
 Db 31 DTYIH 35

RESULT 40  
 AAR47040  
 ID AAR47040 standard; peptide; 120 AA.  
 XX AAR47040;  
 AC AAR47040;

DT 25-MAR-2003 (revised)  
 DT 02-SEP-1994 (first entry)

XX Sequence of the heavy chain variable domain of a humanised version of  
 DE muMAB4D5.  
 DE Monoclonal antibody; heavy chain; variable domain; humanised.

XX Synthetic.  
 OS WO9404679-A1.  
 PN 03-MAR-1994.  
 PD 20-AUG-1993; 93WO-US007832.  
 XX 21-AUG-1992; 92US-00934373.  
 PR (GETH ) GENENTECH INC.  
 PA Carter PJ, Presta LG;  
 PI WPI; 1994-083196/10.  
 DR Preparation of improved humanised antibodies - by comparison of consensus  
 XX and import complementarity determining regions and framework region  
 PT sequences, e.g. to humanise murine.  
 XX Claim 14; Page 108; 126pp; English.  
 PS The murine antibody muMAB4D5 is directed against the extracellular domain  
 CC (ECD) of p185-HER2. The humanised antibody is capable of binding to a  
 CC predetermined antigen. It comprises a FR region having substantially the

CC AA sequence of a human immunoglobulin and a CDR having substantially the  
 CC AA sequence of a non-human immunoglobulin. These novel anti-p185-HER  
 CC antibodies are referred to a huMAb4D5. (Updated on 25-MAR-2003 to correct  
 CC PN field.)  
 XX

SQ Sequence 120 AA;

Query Match 96.7%; Score 29; DB 2; Length 120;

Best Local Similarity 80.0%; Pred. No. 2e+02; 0; Indels 0; Gaps 0;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYYH 5

Db 31 DTYYH 35

Search completed: June 3, 2005, 15:56:42  
 Job time : 48.0606 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 15:49:21 ; Search time 11.6667 Seconds  
(without alignments)  
31.992 Million cell updates/sec.

Title: US-10-089-452-21

Perfect score: 30

Sequence: 1 DTYVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTCUS COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	1148	4	US-09-538-092-156
2	29	96.7	5	3	Sequence 156, App
3	29	96.7	101	4	US-08-646-265A-115
4	29	96.7	117	3	US-09-248-796A-21830
5	29	96.7	120	2	US-08-646-265A-132
6	29	96.7	120	2	US-07-934-373C-2
7	29	96.7	120	2	US-07-934-373C-6
8	29	96.7	120	2	US-07-934-373C-42
9	29	96.7	120	2	US-07-934-373C-44
10	29	96.7	120	2	US-07-934-373C-48
11	29	96.7	120	3	US-08-437-642B-2
12	29	96.7	120	3	US-08-437-642B-6
13	29	96.7	120	3	US-08-437-642B-42
14	29	96.7	120	3	US-08-437-642B-44
15	29	96.7	120	4	US-08-146-206C-2
16	29	96.7	120	4	US-08-146-206C-6
17	29	96.7	120	4	US-09-705-686-2
18	29	96.7	120	4	US-09-705-686-6
19	29	96.7	120	4	US-09-705-392A-2
20	29	96.7	120	4	US-09-705-392A-6
21	29	96.7	120	4	US-09-705-392A-28
22	29	96.7	120	4	US-09-705-392A-30
23	29	96.7	120	4	US-09-705-392A-32
24	29	96.7	120	4	US-09-705-392A-34
25	29	96.7	120	4	US-09-705-392A-36
26	29	96.7	120	4	US-09-705-398-2
27	29	96.7	120	4	US-09-705-398-6
					Sequence 39, Appl

```

28      29      96.7      120      5      PCT-US93-07832-2      Sequence 2, Appli
29      29      96.7      120      5      PCT-US93-07832-6      Sequence 6, Appli
30      29      96.7      123      2      US-08-561-521-9      Sequence 9, Appli
31      29      96.7      123      2      US-08-561-521-11     Sequence 11, Appli
32      29      96.7      123      5      PCT-US95-01219-9     Sequence 9, Appli
33      29      96.7      123      5      PCT-US95-01219-11    Sequence 11, Appli
34      29      96.7      136      3      US-08-646-265A-29    Sequence 29, Appli
35      29      96.7      136      3      US-08-646-265A-99    Sequence 99, Appli
36      29      96.7      140      2      US-08-561-521-4      Sequence 4, Appli
37      29      96.7      140      5      PCT-US95-01219-4     Sequence 4, Appli
38      29      96.7      142      2      US-08-561-521-17     Sequence 17, Appli
39      29      96.7      142      5      PCT-US95-01219-17    Sequence 17, Appli
40      29      96.7      162      4      US-09-710-279-1010   Sequence 1010, Ap
41      29      96.7      162      4      US-09-710-279-2432   Sequence 2432, Ap
42      29      96.7      218      4      US-09-543-681A-4284   Sequence 4284, Ap
43      29      96.7      263      4      US-09-710-279-2750   Sequence 2750, Ap
44      29      96.7      269      3      US-08-646-265A-109   Sequence 109, App
45      29      96.7      292      3      US-09-134-001C-3479   Sequence 3479, Ap

```

#### ALIGNMENTS

RESULT 1

US-09-538-092-156

; Sequence 156, Application US/09538092

; Patent No. 6753314

; GENERAL INFORMATION:

; APPLICANT: Giot, Loic

; APPLICANT: Mansfield, Traci A.

; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same

; FILE REFERENCE: 15966-542

; CURRENT APPLICATION NUMBER: US/09/538,092

; CURRENT FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: 60/127,352

; PRIOR FILING DATE: 1999-04-01

; PRIOR APPLICATION NUMBER: 60/178,965

; PRIOR FILING DATE: 2000-02-01

; NUMBER OF SEQ ID NOS: 1387

; SOFTWARE: CuratSeqFormatter Version 0.9

; SEQ ID NO 156

; LENGTH: 1148

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (0)..(0)

; OTHER INFORMATION: Polypeptide Accession Number YDR128W

US-09-538-092-156

Query Match 100.0%; Score 30; DB 4; Length 1148;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5

Db 136 DTYVH 140

RESULT 2

US-08-646-265A-115

; Sequence 115, Application US/08646265A

; Patent No. 6214973

; GENERAL INFORMATION:

; APPLICANT: OHTOMO, Toshihiko

; APPLICANT: SATO, Koh

; APPLICANT: TSUCHIYA, Masayuki

; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN

; TITLE OF INVENTION: MEDULLOBLASTOMA CELLS

; NUMBER OF SEQUENCES: 132

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

```
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,265A
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01763
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-291078
; FILING DATE: 19-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-646-265A-115

Query Match          96.7%; Score 29; DB 3; Length 5;
Best Local Similarity 80.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5
Db 1 DTVIH 5

RESULT 3
US-09-248-796A-21830
; Sequence 21830, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 21830
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-21830

Query Match          96.7%; Score 29; DB 4; Length 101;
Best Local Similarity 80.0%; Pred. No. 43;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5
Db 58 DTVIH 62

; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,265A
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01763
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-291078
; FILING DATE: 19-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-646-265A-115

Query Match          96.7%; Score 29; DB 3; Length 5;
Best Local Similarity 80.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5
Db 1 DTVIH 5

RESULT 4
US-08-646-265A-132
; Sequence 132, Application US/08646265A
; Patent No. 6214973
; GENERAL INFORMATION:
; APPLICANT: OHTOMO, Toshihiko
; APPLICANT: SATO, Koh
; APPLICANT: TSUCHIYA, Masayuki
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
; TITLE OF INVENTION: MEDULLOBLASTOMA CELLS
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,265A
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01763
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-291078
; FILING DATE: 19-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-646-265A-132

Query Match          96.7%; Score 29; DB 3; Length 117;
Best Local Similarity 80.0%; Pred. No. 49;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5
Db 31 DTVIH 35

RESULT 5
US-07-934-373C-2
; Sequence 2, Application US/07934373C
; Patent No. 5821337
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
```

STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/934,373C  
FILING DATE: 21-Aug-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/05126  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/715272  
FILING DATE: 14-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0709P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-1994  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-07-934-373C-2

Query Match 96.7%; Score 29; DB 2; Length 120;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||:|  
Db 31 DTYIH 35

RESULT 6  
US-07-934-373C-6  
Sequence 6, Application US/07934373C  
Patent No. 5821337  
GENERAL INFORMATION:  
APPLICANT: Paul J. Carter  
APPLICANT: Leonard G. Presta  
TITLE OF INVENTION: Immunglobulin Variants  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/934,373C  
FILING DATE: 21-Aug-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/05126  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/715272  
FILING DATE: 14-JUN-1991  
ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0709P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-1994  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-07-934-373C-6

Query Match 96.7%; Score 29; DB 2; Length 120;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||:|  
Db 31 DTYIH 35

RESULT 7  
US-07-934-373C-42  
Sequence 42, Application US/07934373C  
Patent No. 5821337  
GENERAL INFORMATION:  
APPLICANT: Paul J. Carter  
APPLICANT: Leonard G. Presta  
TITLE OF INVENTION: Immunglobulin Variants  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/934,373C  
FILING DATE: 21-Aug-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/05126  
FILING DATE: 15-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/715272  
FILING DATE: 14-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0709P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-1994  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-07-934-373C-42

Query Match 96.7%; Score 29; DB 2; Length 120;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||:|

```
Db          31 DTYIH 35

RESULT 8
US-07-934-373C-44
; Sequence 44, Application US/07934373C
; Patent No. 5821337
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,373C
; FILING DATE: 21-Aug-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-07-934-373C-44

Query Match          96.7%; Score 29; DB 2; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy          1 DTVVH 5
Db          31 DTYIH 35

RESULT 9
US-07-934-373C-48
; Sequence 48, Application US/07934373C
; Patent No. 5821337
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,642B
; FILING DATE: 09-May-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:

Query Match          96.7%; Score 29; DB 2; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy          1 DTVVH 5
Db          31 DTYIH 35

RESULT 10
US-08-437-642B-2
; Sequence 2, Application US/08437642B
; Patent No. 6054297
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,642B
; FILING DATE: 09-May-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-437-642B-2
Query Match          96.7%; Score 29; DB 3; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DTYVH 5
      |||:|
Db      31 DTYIH 35

RESULT 11
US-08-437-642B-6
; Sequence 6, Application US/08437642B
; Patent No. 6054297
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,642B
; FILING DATE: 09-May-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
```

```
; TOPOLOGY: Linear
US-08-437-642B-6
Query Match          96.7%; Score 29; DB 3; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DTYVH 5
      |||:|
Db      31 DTYIH 35

RESULT 12
US-08-437-642B-42
; Sequence 42, Application US/08437642B
; Patent No. 6054297
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,642B
; FILING DATE: 09-May-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-437-642B-42
Query Match          96.7%; Score 29; DB 3; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DTYVH 5
      |||:|
Db      31 DTYIH 35

RESULT 13
```

US-08-437-642B-44  
; Sequence 44, Application US/08437642B  
; Patent No. 6054297  
; GENERAL INFORMATION:  
; APPLICANT: Paul J. Carter  
; APPLICANT: Leonard G. Presta  
; TITLE OF INVENTION: Immunoglobulin Variants  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/437,642B  
; FILING DATE: 09-May-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/934373  
; FILING DATE: 21-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/146206  
; FILING DATE: 17-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/05126  
; FILING DATE: 15-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P2C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 120 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-437-642B-44

Query Match 96.7%; Score 29; DB 3; Length 120;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 DTYVH 5  
Db 31 DTYIH 35

RESULT 14  
US-08-146-206C-2  
; Sequence 2, Application US/08146206C  
; Patent No. 6407213  
; GENERAL INFORMATION:  
; APPLICANT: Carter, Paul J.  
; APPLICANT: Presta, Leonard G.  
; TITLE OF INVENTION: Method for Making Humanized Antibodies  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California

COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/146,206C  
FILING DATE: 17-No. 6407213-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/715272  
FILING DATE: 14-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0709P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-1994  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-146-206C-2  
  
Query Match 96.7%; Score 29; DB 4; Length 120;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 DTYVH 5  
Db 31 DTYIH 35  
  
RESULT 15  
US-08-146-206C-6  
; Sequence 6, Application US/08146206C  
; Patent No. 6407213  
; GENERAL INFORMATION:  
; APPLICANT: Carter, Paul J.  
; APPLICANT: Presta, Leonard G.  
; TITLE OF INVENTION: Method for Making Humanized Antibodies  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/146,206C  
; FILING DATE: 17-No. 6407213-1993  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 6:  
; INFORMATION FOR SEQ ID NO: 6:

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-146-206C-6
Query Match 96.7%; Score 29; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 31 DTYIH 35

RESULT 16
US-09-705-686-2
; Sequence 2, Application US/09705686
; Patent No. 6639055
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; Presta, Leonard G.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/705.686
; FILING DATE: 02-No. 6639055-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P1D3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-705-686-2
Query Match 96.7%; Score 29; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 31 DTYIH 35

RESULT 17
US-09-705-686-6
; Sequence 6, Application US/09705686
; Patent No. 6639055
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; Presta, Leonard G.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/705.392A
; FILING DATE: 02-No. 6639055-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P1D3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-705-686-6
Query Match 96.7%; Score 29; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 31 DTYIH 35

RESULT 18
US-09-705-392A-2
; Sequence 2, Application US/09705392A
; Patent No. 6719971
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; Presta, Leonard G.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/705.392A
; FILING DATE: 02-No. 6639055-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P1D3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-705-686-6
Query Match 96.7%; Score 29; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 31 DTYIH 35
```

```
; FILING DATE: 02-NO. 6719971-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P1D1 REVISED
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-705-392A-2
Query Match          96.7%; Score 29; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DTYVH 5
Db      31 DTYIH 35.

RESULT 19
US-09-705-392A-6
; Sequence 6, Application US/09705392A
; Patent No. 6719971
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; Presta, Leonard G.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09705,392A
; FILING DATE: 02-NO. 6719971-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P1D1 REVISED
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
```

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-705-392A-6
Query Match          96.7%; Score 29; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DTYVH 5
Db      31 DTYIH 35.

RESULT 20
US-09-705-392A-28
; Sequence 28, Application US/09705392A
; Patent No. 6719971
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; Presta, Leonard G.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09705,392A
; FILING DATE: 02-NO. 6719971-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P1D1 REVISED
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-705-392A-28
Query Match          96.7%; Score 29; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DTYVH 5
Db      31 DTYIH 35.

RESULT 21
US-09-705-392A-30
; Sequence 30, Application US/09705392A
; Patent No. 6719971
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; Presta, Leonard G.
```



;  
; TITLE OF INVENTION: Method for Making Humanized Antibodies  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
;  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
;  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/705,392A  
; FILING DATE: 02-NOV-1993  
; CLASSIFICATION: <Unknown>  
;  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/146206  
; FILING DATE: 17-NOV-1993  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P1D1 REVISED  
;  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
;  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 120 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
;  
; SEQUENCE DESCRIPTION: SEQ ID NO: 30:  
US-09-705-392A-30  
  
Query Match 96.7%; Score 29; DB 4; Length 120;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 DTVVH 5  
Db 31 DTYYH 35  
  
RESULT 22  
US-09-705-392A-32  
; Sequence 32, Application US/09705392A  
; Patent No. 6719971  
; GENERAL INFORMATION:  
; APPLICANT: Carter, Paul J.  
; Presta, Leonard G.  
; TITLE OF INVENTION: Method for Making Humanized Antibodies  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
;  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
;  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/705,392A  
; FILING DATE: 02-NOV-1993  
; CLASSIFICATION: <Unknown>  
;  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/146206  
; FILING DATE: 17-NOV-1993  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P1D1 REVISED  
;  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
;  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 120 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
;  
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:  
US-09-705-392A-34

;  
; APPLICATION NUMBER: 08/146206  
; FILING DATE: 17-NOV-1993  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P1D1 REVISED  
;  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
;  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 120 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
;  
; SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
US-09-705-392A-32  
  
Query Match 96.7%; Score 29; DB 4; Length 120;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 DTVVH 5  
Db 31 DTYYH 35  
  
RESULT 23  
US-09-705-392A-34  
; Sequence 34, Application US/09705392A  
; Patent No. 6719971  
; GENERAL INFORMATION:  
; APPLICANT: Carter, Paul J.  
; Presta, Leonard G.  
; TITLE OF INVENTION: Method for Making Humanized Antibodies  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
;  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
;  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/705,392A  
; FILING DATE: 02-NOV-1993  
; CLASSIFICATION: <Unknown>  
;  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/146206  
; FILING DATE: 17-NOV-1993  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P1D1 REVISED  
;  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
;  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 120 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
;  
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:  
US-09-705-392A-34

```
Query Match      96.7%; Score 29; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DTYVH 5
Db      31 DTYIH 35

RESULT 24
US-09-705-392A-36
; Sequence 36, Application US/09705392A
; Patent No. 6719971
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; Presta, Leonard G.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/705,392A
; FILING DATE: 02-Nov. 6719971-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P1D1 REVISED
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-705-392A-36

Query Match      96.7%; Score 29; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DTYVH 5
Db      31 DTYIH 35

RESULT 25
US-09-705-398-2
; Sequence 2, Application US/09705398
; Patent No. 6800738
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; Presta, Leonard G.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/705,398
; FILING DATE: 02-Nov. 6800738-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P1D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-705-398-2

Query Match      96.7%; Score 29; DB 4; Length 120;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 DTYVH 5
Db      31 DTYIH 35

RESULT 26
US-09-705-398-6
; Sequence 6, Application US/09705398
; Patent No. 6800738
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; Presta, Leonard G.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/705,398
; FILING DATE: 02-Nov. 6800738-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
```

; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709PID2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 120 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-705-398-6  
Query Match 96.7%; Score 29; DB 4; Length 120;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DTYVH 5  
Db 31 DTYIH 35  
RESULT 27  
US-09-232-290-39  
; Sequence 39, Application US/09232290A  
; Patent No. 6815540  
; GENERAL INFORMATION:  
; APPLICANT: PLUCKTHUN, ANDREAS  
; APPLICANT: NIERA, LARS  
; APPLICANT: HONEGGER, ANNEMARIE  
; TITLE OF INVENTION: IMMUNOGLOBULIN SUPER FAMILY DOMAINS AND FRAGMENTS WITH  
; FILE REFERENCE: MORPHO/7  
; CURRENT FILING DATE: 1999-01-15  
; EARLIER APPLICATION NUMBER: PCT/EP96/02230  
; EARLIER FILING DATE: 1996-05-23  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 39  
; LENGTH: 120  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Humanized  
US-09-232-290-39  
Query Match 96.7%; Score 29; DB 4; Length 120;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DTYVH 5  
Db 31 DTYIH 35  
RESULT 28  
PCT-US93-07832-2  
; Sequence 2, Application PC/TUS9307832  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; TITLE OF INVENTION: Immunoglobulin Variants  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA

; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/07832  
; FILING DATE: 19930820  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/05126  
; FILING DATE: 15-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/934373  
; FILING DATE: 21-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME:  
; REGISTRATION NUMBER:  
; REFERENCE/DOCKET NUMBER: 709P2PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE:  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 120 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; PCT-US93-07832-2  
Query Match 96.7%; Score 29; DB 5; Length 120;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DTYVH 5  
Db 31 DTYIH 35  
RESULT 29  
PCT-US93-07832-6  
; Sequence 6, Application PC/TUS9307832  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; TITLE OF INVENTION: Immunoglobulin Variants  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/07832  
; FILING DATE: 19930820  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/05126  
; FILING DATE: 15-JUN-1992  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/934373  
FILING DATE: 21-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME:  
REGISTRATION NUMBER:  
REFERENCE/DOCKET NUMBER: 709P2PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE:  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US93-07832-6

Query Match 96.7%; Score 29; DB 5; Length 120;  
Best Local Similarity 80.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||:|  
Db 31 DTYIH 35

RESULT 30  
US-08-561-521-9  
; Sequence 9, Application US/08561521  
; Patent No. 5840299  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.  
; APPLICANT: Saldanha, Jose  
; APPLICANT: Jones, S. Tarran  
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/561,521  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/186,269A  
; FILING DATE: 25-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William L.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15270-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 123 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-561-521-9

Query Match 96.7%; Score 29; DB 2; Length 123;  
Best Local Similarity 80.0%; Pred. No. 52;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||:|  
Db 31 DTYIH 35

RESULT 31  
US-08-561-521-11  
; Sequence 11, Application US/08561521  
; Patent No. 5840299  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.  
; APPLICANT: Saldanha, Jose  
; APPLICANT: Jones, S. Tarran  
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/561,521  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/186,269A  
; FILING DATE: 25-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William L.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15270-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 123 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-561-521-11

Query Match 96.7%; Score 29; DB 2; Length 123;  
Best Local Similarity 80.0%; Pred. No. 52;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||:|  
Db 31 DTYIH 35

RESULT 32  
PCT-US95-01219-9  
; Sequence 9, Application PC/TUS9501219  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.  
; APPLICANT: Saldanha, Jose  
; APPLICANT: Jones, S. Tarran

```
;
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
;
; TITLE OF INVENTION: Adhesion Molecule VIA-4
;
; NUMBER OF SEQUENCES: 45
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Townsend and Townsend Kourie and Crew
;
; STREET: One Market Plaza, Steuart Tower, Suite 2000
;
; CITY: San Francisco
;
; STATE: California
;
; COUNTRY: USA
;
; ZIP: 94105
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: PCT/US95/01219
;
; FILING DATE: 25-JAN-1995
;
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER: US 08/186,269
;
; FILING DATE: 25-JAN-1994
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Smith, William L.
;
; REGISTRATION NUMBER: 30,223
;
; REFERENCE/DOCKET NUMBER: 15270-14
;
; TELEPHONE: 415-543-9600
;
; TELEFAX: 415-543-5043
;
; INFORMATION FOR SEQ ID NO: 9:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 123 amino acids
;
; TYPE: amino acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; PCT-US95-01219-9
;
; Query Match 96.7%; Score 29; DB 5; Length 123;
; Best Local Similarity 80.0%; Pred. No. 52;
; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 DTVVH 5
;
; Db 31 DTYIH 35
;
; RESULT 33
;
; PCT-US95-01219-11
;
; Sequence 11, Application PC/TUS9501219
;
; GENERAL INFORMATION:
;
; APPLICANT: Bendig, Mary M.
;
; APPLICANT: Leger, Olivier J.
;
; APPLICANT: Saldanha, Jose
;
; APPLICANT: Jones, S. Tarran
;
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
;
; NUMBER OF SEQUENCES: 45
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Townsend and Townsend Kourie and Crew
;
; STREET: One Market Plaza, Steuart Tower, Suite 2000
;
; CITY: San Francisco
;
; STATE: California
;
; COUNTRY: USA
;
; ZIP: 94105
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: PCT/US95/01219
;
; FILING DATE: 25-JAN-1995
```

```
;
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER: US 08/186,269
;
; FILING DATE: 25-JAN-1994
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Smith, William L.
;
; REGISTRATION NUMBER: 30,223
;
; REFERENCE/DOCKET NUMBER: 15270-14
;
; TELEPHONE: 415-543-9600
;
; TELEFAX: 415-543-5043
;
; INFORMATION FOR SEQ ID NO: 11:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 123 amino acids
;
; TYPE: amino acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; PCT-US95-01219-11
;
; Query Match 96.7%; Score 29; DB 5; Length 123;
; Best Local Similarity 80.0%; Pred. No. 52;
; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 DTVVH 5
;
; Db 31 DTYIH 35
;
; RESULT 34
;
; US-08-646-265A-29
;
; Sequence 29, Application US/08646265A
;
; Patent No. 6214973
;
; GENERAL INFORMATION:
;
; APPLICANT: OHTOMO, Toshihiko
;
; APPLICANT: SATO, Koh
;
; APPLICANT: TSUCHIYA, Masayuki
;
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
;
; NUMBER OF SEQUENCES: 132
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Foley & Lardner
;
; STREET: 3000 K Street, N.W., Suite 500
;
; CITY: Washington
;
; STATE: D.C.
;
; COUNTRY: USA
;
; ZIP: 20007-5109
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/646,265A
;
; FILING DATE: 09-SEP-1996
;
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER: WO PCT/JP94/01763
;
; FILING DATE: 19-OCT-1994
;
; APPLICATION NUMBER: JP 5-291078
;
; FILING DATE: 19-NOV-1993
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: WEGNER, Harold C.
;
; REGISTRATION NUMBER: 25,258
;
; REFERENCE/DOCKET NUMBER: 53466/184
;
; TELEPHONE: (202)672-5300
;
; TELEFAX: (202)672-5399
;
; TELEX: 904136
;
; INFORMATION FOR SEQ ID NO: 29:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 136 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-646-265A-29

Query Match          96.7%; Score 29; DB 3; Length 136;
Best Local Similarity 80.0%; Pred. No. 57;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 50 DTYIH 54

RESULT 35
US-08-646-265A-99
; Sequence 99, Application US/08646265A
; Patent No. 6214973
; GENERAL INFORMATION:
; APPLICANT: OHTOMO, Toshiniko
; APPLICANT: SATO, Koh
; APPLICANT: TSUCHIYA, Masayuki
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
; TITLE OF INVENTION: MEDULLOBLASTOMA CELLS
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,265A
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01763
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-291078
; FILING DATE: 19-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 136 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-646-265A-99

Query Match          96.7%; Score 29; DB 3; Length 136;
Best Local Similarity 80.0%; Pred. No. 57;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 50 DTYIH 54

RESULT 36
US-08-561-521-4
; Sequence 4, Application US/08561521
; Patent No. 5840299
; GENERAL INFORMATION:
; APPLICANT: Bendig, Mary M.
; APPLICANT: Leger, Olivier J.
; APPLICANT: Saldanha, Jose
; APPLICANT: Jones, S. Tarran
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
; TITLE OF INVENTION: Adhesion Molecule VLA-4
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/561,521
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,269A
; FILING DATE: 25-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William L.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 15270-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 140 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-561-521-4

Query Match          96.7%; Score 29; DB 2; Length 140;
Best Local Similarity 80.0%; Pred. No. 59;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 50 DTYIH 54

RESULT 37
PCT-US95-01219-4
; Sequence 4, Application PC/TUS9501219
; GENERAL INFORMATION:
; APPLICANT: Bendig, Mary M.
; APPLICANT: Leger, Olivier J.
; APPLICANT: Saldanha, Jose
; APPLICANT: Jones, S. Tarran
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
; TITLE OF INVENTION: Adhesion Molecule VLA-4
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01219  
FILING DATE: 25-JAN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/186,269  
FILING DATE: 25-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William L.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 15270-14  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 140 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-01219-4

Query Match 96.7%; Score 29; DB 5; Length 140;  
Best Local Similarity 80.0%; Pred. No. 59;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 50 DTYIH 54

RESULT 38  
US-08-561-521-17  
Sequence 17, Application US/08561521  
Patent No. 5840299  
GENERAL INFORMATION:  
APPLICANT: Bendig, Mary M.  
APPLICANT: Leger, Olivier J.  
APPLICANT: Saldanha, Jose  
APPLICANT: Jones, S. Tarran  
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/561,521  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/186,269A  
FILING DATE: 25-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William L.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 15270-14  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043

INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 142 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-561-521-17

Query Match 96.7%; Score 29; DB 2; Length 142;  
Best Local Similarity 80.0%; Pred. No. 59;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 50 DTYIH 54

RESULT 39  
PCT-US95-01219-17  
Sequence 17, Application PC/TUS9501219  
GENERAL INFORMATION:  
APPLICANT: Bendig, Mary M.  
APPLICANT: Leger, Olivier J.  
APPLICANT: Saldanha, Jose  
APPLICANT: Jones, S. Tarran  
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01219  
FILING DATE: 25-JAN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/186,269  
FILING DATE: 25-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William L.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 15270-14  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 142 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-01219-17

Query Match 96.7%; Score 29; DB 5; Length 142;  
Best Local Similarity 80.0%; Pred. No. 59;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 50 DTYIH 54

RESULT 40  
US-09-710-279-1010

; Sequence 1010, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1010  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: amino acid sequence  
US-09-710-279-1010

Query Match 96.7%; Score 29; DB 4; Length 162;  
Best Local Similarity 80.0%; Pred. No. 68;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYYH 5  
|||:  
Db 27 DTYYH 31

Search completed: June 3, 2005, 16:03:55  
Job time : 12.6667 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 3, 2005, 16:01:46 ; Search time 36.2121 Seconds  
(without alignments)  
47.730 Million cell updates/sec

Title: US-10-089-452-21

Perfect score: 30

Sequence: 1 DTYVH 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1465611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 1465611

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	35	15 US-10-767-701-52529	Sequence 52529, A
2	30	100.0	42	15 US-10-424-599-218006	Sequence 218006, A
3	30	100.0	71	9 US-09-867-550-1586	Sequence 1586, Ap
4	30	100.0	74	15 US-10-424-599-227413	Sequence 227413, A
5	30	100.0	77	15 US-10-424-599-207224	Sequence 207224, A
6	30	100.0	306	16 US-10-755-889-434	Sequence 434, App
7	30	100.0	316	15 US-10-236-115-1377	Sequence 1377, Ap
8	30	100.0	336	15 US-10-425-114-44748	Sequence 44748, A
9	30	100.0	440	15 US-10-094-749-1912	Sequence 1912, Ap
10	30	100.0	491	15 US-10-425-114-45603	Sequence 45603, A
11	30	100.0	498	15 US-10-425-114-49464	Sequence 49464, A
12	30	100.0	563	15 US-10-424-599-216212	Sequence 216212, A
13	30	100.0	600	14 US-10-128-714-8221	Sequence 8221, Ap

14	30	100.0	649	15 US-10-424-599-261273	Sequence 261273, A
15	30	100.0	677	15 US-10-425-114-45922	Sequence 45922, A
16	30	100.0	777	14 US-10-128-714-8221	Sequence 8221, Ap
17	29	96.7	5	10 US-09-749-873-115	Sequence 115, Appl
18	29	96.7	10	15 US-10-410-894-22	Sequence 22, Appl
19	29	96.7	43	15 US-10-424-599-253506	Sequence 253506, A
20	29	96.7	55	16 US-10-437-963-172343	Sequence 172343, A
21	29	96.7	58	15 US-10-425-114-59607	Sequence 59607, A
22	29	96.7	74	16 US-10-437-963-125220	Sequence 125220, A
23	29	96.7	80	15 US-10-424-599-226325	Sequence 226325, A
24	29	96.7	110	14 US-10-179-851-2	Sequence 2, Appli
25	29	96.7	110	14 US-10-179-851-3	Sequence 3, Appli
26	29	96.7	116	10 US-09-910-483-1	Sequence 1, Appli
27	29	96.7	116	10 US-09-910-483-5	Sequence 5, Appli
28	29	96.7	116	10 US-09-910-483-9	Sequence 9, Appli
29	29	96.7	116	10 US-09-910-483-13	Sequence 13, Appl
30	29	96.7	116	10 US-09-910-483-17	Sequence 17, Appl
31	29	96.7	116	10 US-09-910-483-21	Sequence 21, Appl
32	29	96.7	116	10 US-09-910-483-25	Sequence 25, Appl
33	29	96.7	116	10 US-09-910-483-29	Sequence 29, Appl
34	29	96.7	116	10 US-09-910-483-33	Sequence 33, Appl
35	29	96.7	116	10 US-09-910-483-37	Sequence 37, Appl
36	29	96.7	116	10 US-09-910-483-41	Sequence 41, Appl
37	29	96.7	116	10 US-09-910-483-43	Sequence 43, Appl
38	29	96.7	117	10 US-09-749-873-132	Sequence 132, App
39	29	96.7	117	16 US-10-379-392-41	Sequence 41, Appl
40	29	96.7	117	16 US-10-379-392-128	Sequence 128, App
41	29	96.7	117	16 US-10-379-392-132	Sequence 132, App
42	29	96.7	117	16 US-10-379-392-134	Sequence 134, App
43	29	96.7	117	16 US-10-379-392-150	Sequence 150, App
44	29	96.7	117	16 US-10-379-392-181	Sequence 181, App
45	29	96.7	117	16 US-10-379-392-183	Sequence 183, App

## ALIGNMENTS

## RESULT 1

```
US-10-767-701-52529
; Sequence 52529, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 52529
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 12510479.pcp
US-10-767-701-52529
```

Query Match 100.0%; Score 30; DB 16; Length 35;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	DTYVH	5
Db	23	DTYVH	27

## RESULT 2

```
US-10-424-599-218006
; Sequence 218006, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
```

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 218006  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_38887C.1.pep  
US-10-424-599-218006

Query Match 100.0%; Score 30; DB 15; Length 42;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 12 DTYVH 16

## RESULT 3

US-09-867-550-1586  
; Sequence 1586, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1586  
; LENGTH: 71  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-867-550-1586

Query Match 100.0%; Score 30; DB 9; Length 71;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 20 DTYVH 24

## RESULT 4

US-10-424-599-227413  
; Sequence 227413, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 227413  
; LENGTH: 74  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_47383C.1.pep  
US-10-424-599-227413

Query Match 100.0%; Score 30; DB 15; Length 74;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 36 DTYVH 40

## RESULT 5

US-10-424-599-207224  
; Sequence 207224, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 207224  
; LENGTH: 77  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_29149C.1.pep  
US-10-424-599-207224

Query Match 100.0%; Score 30; DB 15; Length 77;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 60 DTYVH 64

## RESULT 6

US-10-755-889-434  
; Sequence 434, Application US/10755889  
; Publication No. US20040171823A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB  
; TITLE OF INVENTION: PATHWAY  
; FILE REFERENCE: D0284 NP  
; CURRENT APPLICATION NUMBER: US/10/755,889  
; CURRENT FILING DATE: 2004-01-13  
; PRIOR APPLICATION NUMBER: U.S. 60/440,068  
; PRIOR FILING DATE: 2003-01-14  
; PRIOR APPLICATION NUMBER: U.S. 60/469,757  
; PRIOR FILING DATE: 2003-05-12  
; NUMBER OF SEQ ID NOS: 823  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 434  
; LENGTH: 306  
; TYPE: PRT  
; ORGANISM: Homo sapiens

Query Match 100.0%; Score 30; DB 9; Length 71;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 20 DTYVH 24

US-10-755-889-434

Query Match 100.0%; Score 30; DB 16; Length 306;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||||  
Db 215 DTYVH 219

RESULT 7

US-10-296-115-1377  
; Sequence 1377, Application US/10296115  
; Publication No. US20040053248A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq Inc  
; TITLE OF INVENTION: NO. US20040053248A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 784PCT  
; CURRENT APPLICATION NUMBER: US/10/296,115  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: US09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 1478  
; SEQ ID NO 1377  
; LENGTH: 316  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-296-115-1377

Query Match 100.0%; Score 30; DB 15; Length 316;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||||  
Db 225 DTYVH 229

RESULT 8

US-10-425-114-44748  
; Sequence 44748, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 44748  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 701039608\_FLI.pap  
US-10-425-114-44748

Query Match 100.0%; Score 30; DB 15; Length 336;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||||  
Db 14 DTYVH 18

RESULT 9

US-10-094-749-1912  
; Sequence 1912, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, KYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1912  
; LENGTH: 440  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-094-749-1912

Query Match 100.0%; Score 30; DB 15; Length 440;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||||  
Db 257 DTYVH 261

RESULT 10

US-10-425-114-45603  
; Sequence 45603, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 45603  
; LENGTH: 491  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700651974\_FLI.pap

US-10-425-114-45603

Query Match 100.0%; Score 30; DB 15; Length 491;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5

Db 169 DTYVH 173

RESULT 11

US-10-425-114-49464  
; Sequence 49464, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 49464  
; LENGTH: 498  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700651482\_FLI.pep

US-10-425-114-49464

Query Match 100.0%; Score 30; DB 15; Length 498;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5

Db 176 DTYVH 180

RESULT 12

US-10-424-599-216212  
; Sequence 216212, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 216212  
; LENGTH: 563  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_37269C.1.pep

US-10-424-599-216212

Query Match 100.0%; Score 30; DB 15; Length 563;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5

Db 503 DTYVH 507

RESULT 13

US-10-128-714-3221  
; Sequence 3221, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Mengdi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Broshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 3221  
; LENGTH: 600  
; TYPE: PRT  
; ORGANISM: Aspergillus fumigatus

US-10-128-714-3221

Query Match 100.0%; Score 30; DB 14; Length 600;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5

Db 414 DTYVH 418

RESULT 14

US-10-424-599-261273  
; Sequence 261273, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Cao Yongwei  
; APPLICANT: Zhou Yihua  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 261273  
; LENGTH: 649  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(649)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_77952C.1.pep

US-10-424-599-261273

Query Match 100.0%; Score 30; DB 15; Length 649;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 327 DTYVH 331

## RESULT 15

US-10-425-114-45922  
; Sequence 45922, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 45922  
; LENGTH: 677  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 701011172\_FLI.pap  
US-10-425-114-45922

Query Match 100.0%; Score 30; DB 15; Length 677;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 355 DTYVH 359

## RESULT 16

US-10-128-714-8221  
; Sequence 8221, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; FILE OF INVENTION: Methods of Use  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8221

; LENGTH: 777

; TYPE: PRT  
; ORGANISM: Aspergillus fumigatus  
US-10-128-714-8221

Query Match 100.0%; Score 30; DB 14; Length 777;  
Best Local Similarity 100.0%; Pred. No. 6.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 591 DTYVH 595

## RESULT 17

US-09-749-873-115  
; Sequence 115, Application US/09749873  
; Publication No. US20030023045A1  
; GENERAL INFORMATION:  
; APPLICANT: OHTOMO, Toshihiko  
; APPLICANT: SATO, Koh  
; APPLICANT: TSUCHIYA, Masayuki  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; MEDULLOBLASTOMA CELLS  
; NUMBER OF SEQUENCES: 132  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/749,873  
; FILING DATE: 29-Dec-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/646,265  
; FILING DATE: 1996-09-09  
; APPLICATION NUMBER: JP 5-291078  
; FILING DATE: 19-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WEGNER, Harold C.  
; REGISTRATION NUMBER: 25,258  
; REFERENCE/DOCKET NUMBER: 53466/184  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 115:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 115:  
US-09-749-873-115

Query Match 96.7%; Score 29; DB 10; Length 5;  
Best Local Similarity 80.0%; Pred. No. 1.3e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 1 DTYIH 5

## RESULT 18

US-10-410-894-22  
; Sequence 22, Application US/10410894  
; Publication No. US2003022863A1  
; GENERAL INFORMATION:  
; APPLICANT: GENENTECH, INC.  
; APPLICANT: LOWMAN, Henry B.  
; APPLICANT: GERSTNER, Resi B.  
; APPLICANT: CARTER, Paul J.  
; TITLE OF INVENTION: ANTI-HER2 ANTIBODY VARIANTS  
; FILE REFERENCE: 39766-0108 US  
; CURRENT APPLICATION NUMBER: US/10/410,894  
; CURRENT FILING DATE: 2003-04-09  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-410-894-22

Query Match 96.7%; Score 29; DB 15; Length 10;  
Best Local Similarity 80.0%; Pred. No. 14;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5  
|||:|  
Db 6 DTVIH 10

## RESULT 19

US-10-424-599-253506  
; Sequence 253506, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 253506  
; LENGTH: 43  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_70940C.1.pep  
US-10-424-599-253506

Query Match 96.7%; Score 29; DB 15; Length 43;  
Best Local Similarity 80.0%; Pred. No. 58;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5  
|||:|  
Db 2 DTVIH 6

## RESULT 20

US-10-437-963-172343  
; Sequence 172343, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 172343  
; LENGTH: 55  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_70489C.1.pep  
US-10-437-963-172343

Query Match 96.7%; Score 29; DB 16; Length 55;  
Best Local Similarity 80.0%; Pred. No. 74;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5  
|||:|  
Db 41 DTVIH 45

## RESULT 21

US-10-425-114-59607  
; Sequence 59607, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 59607  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3595-036-B5\_FLI.pep  
US-10-425-114-59607

Query Match 96.7%; Score 29; DB 15; Length 58;  
Best Local Similarity 80.0%; Pred. No. 78;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5  
|||:|  
Db 17 DTVIH 21

## RESULT 22

US-10-437-963-125220  
; Sequence 125220, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 172343  
; LENGTH: 55  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_70489C.1.pep  
US-10-437-963-125220

Query Match 96.7%; Score 29; DB 15; Length 58;  
Best Local Similarity 80.0%; Pred. No. 78;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 125220
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27886C.1.pep
US-10-437-963-125220

Query Match          96.7%; Score 29; DB 16; Length 74;
Best Local Similarity 80.0%; Pred. No. 99;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVVH 5
   |||:|
Db 52 DTVIH 56

RESULT 23
US-10-424-599-226325
; Sequence 226325, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 226325
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_46400C.1.pep
US-10-424-599-226325

Query Match          96.7%; Score 29; DB 15; Length 80;
Best Local Similarity 80.0%; Pred. No. 11e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVVH 5
   |||:|
Db 7 DTVIH 11

RESULT 24
US-10-179-851-2
; Sequence 2, Application US/10179851
; Publication No. US20030017518A1
; GENERAL INFORMATION:
; APPLICANT: Lam et al.
; APPLICANT: Lam et al.
; APPLICANT: Zhao, Xin S
; APPLICANT: Lam, Joseph S
; TITLE OF INVENTION: NON-RADIOACTIVE ASSAY OF LIPOPOLYSACCHARIDE KINASES
; FILE REFERENCE: BEP 3004
; CURRENT APPLICATION NUMBER: US/10/179,851
; CURRENT FILING DATE: 2002-08-20
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Mus musculus
```

```
US-10-179-851-2

Query Match          96.7%; Score 29; DB 14; Length 110;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVVH 5
   |||:|
Db 26 DTVIH 30

RESULT 25
US-10-179-851-3
; Sequence 3, Application US/10179851
; Publication No. US20030017518A1
; GENERAL INFORMATION:
; APPLICANT: Lam et al.
; APPLICANT: Zhao, Xin S
; APPLICANT: Lam, Joseph S
; TITLE OF INVENTION: NON-RADIOACTIVE ASSAY OF LIPOPOLYSACCHARIDE KINASES
; FILE REFERENCE: BEP 3004
; CURRENT APPLICATION NUMBER: US/10/179,851
; CURRENT FILING DATE: 2002-08-20
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(110)
; OTHER INFORMATION:
US-10-179-851-3

Query Match          96.7%; Score 29; DB 14; Length 110;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVVH 5
   |||:|
Db 26 DTVIH 30

RESULT 26
US-09-910-483-1
; Sequence 1, Application US/09910483
; Publication No. US20030035798A1
; GENERAL INFORMATION:
; APPLICANT: FANG, FANG
; APPLICANT: KOHLSTADT, LORI
; APPLICANT: RENO, JOHN
; TITLE OF INVENTION: HUMANIZED ANTIBODIES
; FILE REFERENCE: 014357/027 8772
; CURRENT APPLICATION NUMBER: US/09/910,483
; CURRENT FILING DATE: 2001-07-19
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic VH Domain
; OTHER INFORMATION: peptide of Hum A
US-09-910-483-1

Query Match          96.7%; Score 29; DB 10; Length 116;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

[illegible]



```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic VH Domain
US-09-910-483-21

Query Match          96.7%; Score 29; DB 10; Length 116;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 31 DTYIH 35

RESULT 32
US-09-910-483-25
; Sequence 25, Application US/09910483
; Publication No. US20030035798A1
; GENERAL INFORMATION:
; APPLICANT: FANG, FANG
; APPLICANT: KOHLSTAEDT, LORI
; TITLE OF INVENTION: HUMANIZED ANTIBODIES
; FILE REFERENCE: 014357/027 8772
; CURRENT APPLICATION NUMBER: US/09/910,483
; CURRENT FILING DATE: 2001-07-19
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic VH Domain
US-09-910-483-25

Query Match          96.7%; Score 29; DB 10; Length 116;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 31 DTYIH 35

RESULT 33
US-09-910-483-29
; Sequence 29, Application US/09910483
; Publication No. US20030035798A1
; GENERAL INFORMATION:
; APPLICANT: FANG, FANG
; APPLICANT: KOHLSTAEDT, LORI
; TITLE OF INVENTION: HUMANIZED ANTIBODIES
; FILE REFERENCE: 014357/027 8772
; CURRENT APPLICATION NUMBER: US/09/910,483
; CURRENT FILING DATE: 2001-07-19
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic VH Domain
US-09-910-483-29

Query Match          96.7%; Score 29; DB 10; Length 116;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 31 DTYIH 35

RESULT 34
US-09-910-483-33
; Sequence 33, Application US/09910483
; Publication No. US20030035798A1
; GENERAL INFORMATION:
; APPLICANT: FANG, FANG
; APPLICANT: KOHLSTAEDT, LORI
; TITLE OF INVENTION: HUMANIZED ANTIBODIES
; FILE REFERENCE: 014357/027 8772
; CURRENT APPLICATION NUMBER: US/09/910,483
; CURRENT FILING DATE: 2001-07-19
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic VH Domain
US-09-910-483-33

Query Match          96.7%; Score 29; DB 10; Length 116;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 31 DTYIH 35

RESULT 35
US-09-910-483-37
; Sequence 37, Application US/09910483
; Publication No. US20030035798A1
; GENERAL INFORMATION:
; APPLICANT: FANG, FANG
; APPLICANT: KOHLSTAEDT, LORI
; TITLE OF INVENTION: HUMANIZED ANTIBODIES
; FILE REFERENCE: 014357/027 8772
; CURRENT APPLICATION NUMBER: US/09/910,483
; CURRENT FILING DATE: 2001-07-19
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Murine 1A6 VH Domain
US-09-910-483-37

Query Match          96.7%; Score 29; DB 10; Length 116;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 31 DTYIH 35
```

```
RESULT 36
US-09-910-483-41
; Sequence 41, Application US/09910483
; Publication No. US20030035798A1
; GENERAL INFORMATION:
; APPLICANT: FANG, FANG
; APPLICANT: KOHLSTAEDT, LORI
; TITLE OF INVENTION: HUMANIZED ANTIBODIES
; FILE REFERENCE: 014357/027 8772
; CURRENT APPLICATION NUMBER: US/09/910,483
; CURRENT FILING DATE: 2001-07-19
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Murine 1A6 VH Domain
; OTHER INFORMATION: consensus sequence of Heavy Chain Subgroup III (Humiii)
US-09-910-483-41
Query Match 96.7%; Score 29; DB 10; Length 116;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5
Db 31 DTVIH 35

RESULT 37
US-09-910-483-43
; Sequence 43, Application US/09910483
; Publication No. US20030035798A1
; GENERAL INFORMATION:
; APPLICANT: FANG, FANG
; APPLICANT: KOHLSTAEDT, LORI
; TITLE OF INVENTION: HUMANIZED ANTIBODIES
; FILE REFERENCE: 014357/027 8772
; CURRENT APPLICATION NUMBER: US/09/910,483
; CURRENT FILING DATE: 2001-07-19
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Humanized 1A6
; OTHER INFORMATION: (Hum19) VH Domain consensus sequence of Heavy Chain
; OTHER INFORMATION: Subgroup III (Humiii)
US-09-910-483-43
Query Match 96.7%; Score 29; DB 10; Length 116;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5
Db 31 DTVIH 35

RESULT 38
US-09-749-873-132
; Sequence 132, Application US/09749873
; Publication No. US20030023045A1
; GENERAL INFORMATION:
; APPLICANT: OHTOMO, Toshihiko
; APPLICANT: SATO, Koh

; TSUCHIYA, Masayuki
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
; MEDULLOBLASTOMA CELLS
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/749,873
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/646,265
; FILING DATE: 1996-09-09
; APPLICATION NUMBER: JP 5-291078
; FILING DATE: 19-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 132:
US-09-749-873-132
Query Match 96.7%; Score 29; DB 10; Length 117;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5
Db 31 DTVIH 35

RESULT 39
US-10-379-392-41
; Sequence 41, Application US/10379392
; Publication No. US20040110226A1
; GENERAL INFORMATION:
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John Rudolf
; APPLICANT: Marshall, Shannon Alicia
; APPLICANT: Dahiyat, Basil I.
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION
; FILE REFERENCE: A-71386-3 463077-236
; CURRENT APPLICATION NUMBER: US/10/379,392
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,843
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/384,197
; PRIOR FILING DATE: 2002-05-29
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 41
; LENGTH: 117
; TYPE: PRT
```

```
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Humanized
US-10-379-392-41

Query Match          96.7%; Score 29; DB 16; Length 117;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 31 DTYIH 35

RESULT 40
US-10-379-392-128
; Sequence 128, Application US/10379392
; Publication No. US20040110226A1
; GENERAL INFORMATION:
; APPLICANT: Lazar, Gregory Alan
; APPLICANT: Desjarlais, John Rudolf
; APPLICANT: Marshall, Shannon Alicia
; APPLICANT: Dahiyat, Bassil I.
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION
; FILE REFERENCE: A-71386-3 463077-236
; CURRENT APPLICATION NUMBER: US/10/379,392
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,843
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/384,197
; PRIOR FILING DATE: 2002-05-29
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 128
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-379-392-128

Query Match          96.7%; Score 29; DB 16; Length 117;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 31 DTYIH 35
```

Search completed: June 3, 2005, 16:24:40  
Job time : 37.2121 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 15:48:25 ; Search time 8.78788 Seconds  
(without alignments)  
54.744 Million cell updates/sec

Title: US-10-089-452-21

Perfect score: 30

Sequence: 1 DTYVH 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Pir1.\*
- 2: Pir2.\*
- 3: Pir3.\*
- 4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	30	100.0	207	2 AF0080	probable exported
2	30	100.0	373	2 T34699	hypothetical prote
3	30	100.0	674	2 C44239	amine oxidase (cop
4	30	100.0	695	2 S76751	translation elonga
5	30	100.0	823	2 S18968	cytostatin precurs
6	30	100.0	943	2 S44636	f22b7.5 protein -
7	30	100.0	1148	2 S18555	hypothetical prote
8	29	96.7	85	2 E37262	Ig heavy chain V r
9	29	96.7	587	2 S21139	amine oxidase (cop
10	29	96.7	674	2 JC7251	amine oxidase (cop
11	29	96.7	676	1 WZVZ18	I8 protein - vacci
12	29	96.7	676	2 T37345	NPH-II, helicase -
13	29	96.7	676	2 D42511	I8R protein - vacc
14	29	96.7	676	2 F36843	ATP/GTP-binding pr
15	29	96.7	676	2 T28500	hypothetical prote
16	29	96.7	682	2 D72158	L8R protein - vari
17	29	96.7	713	2 H97827	organic solvent to
18	29	96.7	735	2 T50068	probable ATP-depen
19	29	96.7	770	2 S6805	probable RNA helic
20	29	96.7	1896	2 T08851	down syndrome cell
21	29	96.7	2657	2 T18497	hypothetical prote
22	29	96.7	3119	2 T18414	protein g377 - mal
23	29	96.7	3623	2 T08618	intrinsic factor-B
24	27	90.0	51	2 S36377	Ig heavy chain V r
25	27	90.0	88	2 H46200	retrovirus-related
26	27	90.0	112	2 A49715	Ig kappa chain V r
27	27	90.0	114	4 A47271	nitrophenyl phosph
28	27	90.0	115	2 S03482	Ig heavy chain V-D
29	27	90.0	116	2 S24289	Ig gamma chain V r

30	27	90.0	117	2 S17586	Ig heavy chain V r
31	27	90.0	120	2 S03484	Ig heavy chain V-D
32	27	90.0	120	2 S03471	Ig heavy chain V r
33	27	90.0	122	2 S06823	Ig heavy chain V r
34	27	90.0	123	2 PH1403	Ig heavy chain V r
35	27	90.0	131	2 S52449	Ig kappa chain V r
36	27	90.0	169	2 AG1191	B. subtilis CcpR p
37	27	90.0	178	2 S29594	Ig gamma chain (WM
38	27	90.0	221	2 S49220	Ig gamma-1 chain -
39	27	90.0	268	2 A56446	Ig heavy chain V r
40	27	90.0	308	2 S72886	hypothetical prote
41	27	90.0	328	2 A85959	hydrogenase-2 smal
42	27	90.0	328	2 A80885	hydrogenase-2 smal
43	27	90.0	328	2 B5086	hydrogenase (EC 1.
44	27	90.0	328	2 A31114	hydrogenase-2 smal
45	27	90.0	390	2 F69447	conserved hypothet

ALIGNMENTS

RESULT 1

AF0080  
probable exported protein YPO0651 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AF0080  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AF0080  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-207 <Kur>  
A:Cross-references: UNIPROT:Q8Z163; GB:AL590842; PIDN:CAC89505.1; PID:G15978741; GSPDB:G  
C:Genetics:  
A:Gene: YPO0651  
C:Superfamily: hypothetical protein H1605

Query Match 100.0%; Score 30; DB 2; Length 207;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||  
Db 33 DTYVH 37

RESULT 2

T34699  
hypothetical protein SC1C3.07 SC1C3.07 - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004  
C:Accession: T34699  
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, May 1998  
A:Reference number: Z21554  
A:Accession: T34699  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-373 <OLI>  
A:Cross-references: UNIPROT:O69849; EMBL:AL0233702; PIDN:CAA19231.1; GSPDB:GN00070; SCOE  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC1C3.07

Query Match 100.0%; Score 30; DB 2; Length 373;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 Db 89 DTYVH 93

RESULT 3  
 C44239  
 amine oxidase (copper-containing) (EC 1.4.3.6) precursor - garden pea  
 C:Species: Pisum sativum (garden pea)  
 C>Date: 31-Dec-1993 #sequence\_revision 05-Jan-1996 #text\_change 09-Jul-2004  
 C:Accession: A57327; C44239  
 R:Tipping, A.J.; McPherson, M.J.  
 J. Biol. Chem. 270, 16939-16946, 1995  
 A>Title: Cloning and molecular analysis of the pea seedling copper amine oxidase.  
 A:Reference number: A57327; MUID:95348126; PMID:7622512  
 A:Accession: A57327  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-674 <TIP>  
 A:Cross-references: UNIPROT:Q43077; GB:L39931; NID:G685197; PID:G685198  
 R:James, S.M.; Pačic, M.M.; Scaman, C.H.; Smith, A.J.; Brown, D.E.; Dooley, D.M.; Mure,  
 Biochemistry 31, 12147-12154, 1992  
 A>Title: Identification of topaquinone and its consensus sequence in copper amine oxidase  
 A:Reference number: A44239; MUID:93090748; PMID:1457410  
 A:Accession: C44239  
 A:Molecule type: protein  
 A:Residues: 409-417,'X',419 <JAN>  
 A:Superfamily: amine oxidase (copper-containing)  
 C:Superfamily: amine oxidase (copper-containing)  
 C:Keywords: copper; glycoprotein; metalloprotein; oxidoreductase; quinoprotein; topaquin  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:26-674/Product: amine oxidase (copper-containing) #status predicted <MAT>  
 F:156,389,583/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:382,467,469/Binding site: copper (His) #status predicted  
 F:412/Modified site: topaquinone (Tyr) #status experimental

Query Match 100.0%; Score 30; DB 2; Length 674;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 Db 352 DTYVH 356

RESULT 4  
 S76751  
 translation elongation factor EF-G - Synechocystis sp. (strain PCC 6803)  
 C:Species: Synechocystis sp.  
 A:Variety: PCC 6803  
 C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
 C:Accession: S76751; S47973; S20656  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
 O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
 DNA Res. 3, 109-136, 1996  
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
 S.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S76751  
 A>Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-695 <KAN>  
 A:Cross-references: UNIPROT:P28371; EMBL:D90916; GB:AB001339; NID:G1653715; PIDN:BA01866  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 R:Welsh, P.L.; Johnson, D.R.; Zhang, Y.; Breitenberger, C.A.  
 Plant Mol. Biol. 25, 735-738, 1994  
 A>Title: Synechocystis sp. PCC6803 fuaB gene, located outside of the str operon, encodes  
 A:Reference number: S47973; MUID:94339485; PMID:8061323  
 A:Accession: S47973  
 A:Molecule type: DNA  
 A:Residues: 1-41,'R',43-695 <WE2>  
 A:Cross-references: EMBL:X65159; NID:G46482; PIDN:CAA46277.1; PID:G46483  
 C:Genetics:

A:Gene: fus  
 C:Superfamily: translation elongation factor G; translation elongation factor Tu homolog  
 C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis  
 F:9-136/Domain: translation elongation factor Tu homology <ETU>  
 F:15-22/Region: nucleotide-binding motif A (P-loop)  
 F:133-136/Region: GTP-binding NKXD motif

Query Match 100.0%; Score 30; DB 2; Length 695;  
 Best Local Similarity 100.0%; Pred. No. 83;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 Db 490 DTYVH 494

RESULT 5  
 S18968  
 cyritestin precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
 C:Accession: I48784; S18968  
 R:Senfleben, A.; Wallat, S.; Lemaire, L.; Heinlein, U.A.O.  
 Dev. Growth Differ. 36, 49-58, 1994  
 A>Title: Pre and postmeiotic germ cell specific expression of TAZ83, a gene encoding a p  
 A:Reference number: I48784  
 A:Accession: I48784  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-823 <RES>  
 A:Cross-references: UNIPROT:Q62287; EMBL:X64227; NID:G54264; PID:G54265  
 C:Genetics:  
 A:Gene: TAZ83  
 C:Superfamily: mouse meltrin alpha: disintegrin homology  
 F:395-480/Domain: disintegrin homology <DIS>

Query Match 100.0%; Score 30; DB 2; Length 823;  
 Best Local Similarity 100.0%; Pred. No. 98;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 Db 592 DTYVH 596

RESULT 6  
 S44636  
 f22b7.5 protein - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
 C:Accession: S44636  
 R:Anderson, K.  
 submitted to the EMBL Data Library, March 1993  
 A:Description: Sequence of the C. elegans cosmid F22B7.  
 A:Reference number: S44636  
 A:Accession: S44636  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-943 <AND>  
 A:Cross-references: UNIPROT:P34408; EMBL:L12018; NID:G156298; PID:G156307  
 C:Genetics:  
 F:531-595/Domain: dnaJ amino-terminal homology <DNJ>

Query Match 100.0%; Score 30; DB 2; Length 943;  
 Best Local Similarity 100.0%; Pred. No. 11e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 Db 814 DTYVH 818

## RESULT 7

S51855  
hypothetical protein YDR128w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YD9302.03  
C:Species: Saccharomyces cerevisiae  
C>Date: 05-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 16-Aug-2004  
C:Accession: S51855  
R:Oliver, K.; Harris, D.  
submitted to the EMBL Data Library, February 1995  
A:Reference number: S51853  
A:Accession: S51855  
A:Molecule type: DNA  
A:Residues: 1-1148 <OLI>  
A:Cross-references: UNIPROT:Q03897; EMBL:Z48179; NID:G665660; GSPDB:GN000044  
C:Genetics:  
A:Gene: MIPS:YDR128w  
A:Cross-references: SGD:S0002535  
A:Map position: 4R  
C:Superfamily: WD repeat homology

Query Match 100.0%; Score 30; DB 2; Length 1148;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||||  
Db 136 DTYVH 140

## RESULT 8

E37262  
Ig heavy chain V region (6H2) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 30-Aug-1991 #sequence\_revision 30-Aug-1991 #text\_change 09-May-1997  
C:Accession: E37262  
R:Goshorn, S.C.; Retzel, E.; Jemerson, R.  
J. Biol. Chem. 265, 2134-2142, 1991  
A:Title: Common structural features among monoclonal antibodies binding the same antigen  
A:Reference number: A38601; MUID:91115823; PMID:1703527  
A:Accession: E37262  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-85 <GOS>  
A:Cross-references: GB:M57991  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match 96.7%; Score 29; DB 2; Length 85;  
Best Local Similarity 80.0%; Pred. No. 16;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||||  
Db 9 DTYIH 13

## RESULT 9

S21139  
amine oxidase (copper-containing) (EC 1.4.3.6) precursor - lentil (fragment)  
C:Species: Lens culinaris (lentil)  
C>Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 09-Jul-2004  
C:Accession: S21139; S21229; A48985; B48985  
R:Rossi, A.; Petruzzelli, R.; Agro, A.F.  
FEBS Lett. 301, 253-257, 1992  
A:Title: cDNA-derived amino-acid sequence of lentil seedlings' amine oxidase.  
A:Reference number: S21139; MUID:92249576; PMID:1577161  
A:Accession: S21139  
A:Molecule type: mRNA  
A:Residues: 1-587 <ROS1>  
A:Cross-references: UNIPROT:P49252; EMBL:X64201; NID:g19147; PIDN:CAA45526.1; PID:g19148  
A:Accession: S21229  
A:Molecule type: Protein  
A:Residues: 19-66;542-561 <ROS2>

R:Agro, A.F.; Rossi, A.  
Biochem. Soc. Trans. 20, 369-373, 1992  
A:Title: Copper-containing plant oxidases.  
A:Reference number: A48985; MUID:93012405; PMID:1397633  
A:Accession: A48985  
A:Molecule type: Protein  
A:Residues: 24-35;258-276;360-377 <AGR1>  
A:Experimental source: seedling  
A:Note: sequence extracted from NCBI backbone (NCBIP:116874, NCBIP:116876, NCBIP:116878)  
A:Accession: B48985  
A:Molecule type: mRNA  
A:Residues: 403-408 <AGR2>  
C:Superfamily: amine oxidase (copper-containing)  
C:Keywords: copper; glycoprotein; metalloprotein; oxidoreductase; quinoprotein; topaquin  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-587/Product: amine oxidase (copper-containing) #status experimental <MAT>  
F:149,252,382/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:375,460,462/Binding site: copper (His) #status predicted  
F:405/Modified site: topaquinone (Tyr) #status predicted

Query Match 96.7%; Score 29; DB 2; Length 587;  
Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||||  
Db 345 DTYIH 349

## RESULT 10

JC7251  
amine oxidase (copper-containing) (EC 1.4.3.6) - garden pea  
C:Species: Pium sativum (garden pea)  
C>Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jul-2004  
C:Accession: JC7251  
R:Koyanagi, T.; Matsumura, K.; Kuroda, S.; Tanizawa, K.  
BioSci. Biotechnol. Biochem. 64, 717-722, 2000  
A:Title: Molecular cloning and heterologous expression of pea seedling copper amine oxid  
A:Reference number: JC7251  
A:Accession: JC7251  
A:Molecule type: mRNA  
A:Residues: 1-674 <KOY>  
A:Cross-references: UNIPROT:Q9SKW5; DDBJ:AB026253  
A:Experimental source: seed  
C:Comment: This enzyme, a homodimer, containing one Cu2+ ion and one 2,4,5-trihydroxyph  
diamine substrate and catalyzes the oxidative deamination of mono-, di-, polyamines to  
C:Superfamily: amine oxidase (copper-containing)  
C:Keywords: copper; copper binding; homodimer; oxidoreductase

Query Match 96.7%; Score 29; DB 2; Length 674;  
Best Local Similarity 80.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||||  
Db 352 DTYIH 356

## RESULT 11

WZVZ18  
I8 protein - vaccinia virus (strain WR)  
C:Species: vaccinia virus  
C>Date: 30-Jun-1989 #sequence\_revision 30-Sep-1992 #text\_change 09-Jul-2004  
C:Accession: B38497; H29889  
R:Fathi, Z.; Condit, R.C.  
Virology 181, 258-272, 1991  
A:Title: Genetic and molecular biological characterization of a vaccinia virus temperatu  
A:Reference number: A38497; MUID:91134989; PMID:1994576  
A:Accession: B38497  
A:Molecule type: DNA  
A:Residues: 1-676 <FAT>  
A:Cross-references: UNIPROT:P12927; GB:J03399; EMBL:M29901; NID:g335662; PIDN:AAB59810.1  
R:Schmitt, J.F.C.; Stunnenberg, H.G.

J. Virol. 62, 1889-1897, 1988  
A;Title: Sequence and transcriptional analysis of the vaccinia virus HindIII I fragment.  
A;Reference number: A29889; MUID:88215015; PMID:2835495  
A;Accession: H29889  
A;Molecule type: DNA  
A;Residues: 1-99 <SCH>  
A;Cross-references: GB:J03399  
A;Note: the authors translated the codons TAC, AGT, GAA, CAC, ACA, ATA, and AGC for residues 1-99.  
C;Genetics:  
A;Gene: I8  
C;Superfamily: vaccinia virus I8 protein  
C;Keywords: ATP; nucleotide binding; P-loop  
F;185-192/Region: nucleotide-binding motif A (P-loop)  
F;292-297/Region: nucleotide-binding motif B  
F;296-299/Region: DEXH motif

Query Match 96.7%; Score 29; DB 1; Length 676;  
Best Local Similarity 80.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTYVH 5  
DB 632 DTYIH 636  
|||||

RESULT 12  
T37345  
NPH-II, helicase - vaccinia virus (strain Ankara)  
C;Species: vaccinia virus  
A;Variety: strain Ankara  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C;Accession: T37345  
R;Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dorner, F.  
submitted to the EMBL Data Library, March 1997  
A;Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain  
A;Reference number: 220877  
A;Accession: T37345  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-676 <ANT>  
A;Cross-references: UNIPROT:O57193; EMBL:U94848; PIDN:AAB96491.1  
A;Experimental source: strain Ankara  
C;Genetics:  
A;Note: MVA069R  
C;Superfamily: vaccinia virus I8 protein

Query Match 96.7%; Score 29; DB 2; Length 676;  
Best Local Similarity 80.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTYVH 5  
DB 632 DTYIH 636  
|||||

RESULT 13  
D42511  
I8R protein - vaccinia virus (strain Copenhagen)  
C;Species: vaccinia virus  
A;Note: host Homo sapiens (man)  
C;Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 09-Jul-2004  
C;Accession: D42511  
R;Johnson, G.P.  
submitted to GenBank, June 1990  
A;Reference number: A33172  
A;Accession: D42511  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-676 <JOH>  
A;Cross-references: UNIPROT:P20502  
C;Superfamily: vaccinia virus I8 protein  
C;Keywords: ATP; nucleotide binding; P-loop  
F;185-192/Region: nucleotide-binding motif A (P-loop)

F;292-297/Region: nucleotide-binding motif B  
F;296-299/Region: DEXH motif

Query Match 96.7%; Score 29; DB 2; Length 676;  
Best Local Similarity 80.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTYVH 5  
DB 632 DTYIH 636  
|||||

RESULT 14  
F36843  
ATP/GTP-binding protein K8R - variola virus  
N;Alternate names: K8R protein  
C;Species: variola virus  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C;Accession: F36843; S33077  
R;Blinov, V.M.  
submitted to GenBank, November 1992  
A;Reference number: A36859  
A;Accession: F36843  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-676 <BLI>  
A;Cross-references: UNIPROT:P33051; GB:X69198; NID:G456758; PIDN:CAA49003.1; PID:G297243  
A;Experimental source: strain India-1967, sep. major, isolate Ind3  
R;Shchelkunov, S.N.; Blinov, V.M.; Totmenin, A.V.; Marenikova, S.S.; Kolykhalov, A.A.;  
dzhaparidze, O.G.; Sandakchiev, L.S.  
Virus Res. 27, 25-35, 1993  
A;Title: Nucleotide sequence analysis of variola virus HindIII M, L, I genome fragments.  
A;Reference number: S33069; MUID:93190624; PMID:8383392  
A;Accession: S33077  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-676 <SHC>  
A;Cross-references: EMBL:X67119; NID:G62330; PIDN:CAA47562.1; PID:G62339  
A;Experimental source: strain India-1967, isolate Ind3  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1992  
C;Superfamily: vaccinia virus I8 protein  
C;Keywords: ATP; nucleotide binding; P-loop  
F;185-192/Region: nucleotide-binding motif A (P-loop)  
F;292-297/Region: nucleotide-binding motif B  
F;296-299/Region: DEXH motif

Query Match 96.7%; Score 29; DB 2; Length 676;  
Best Local Similarity 80.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTYVH 5  
DB 632 DTYIH 636  
|||||

RESULT 15  
T28500  
hypothetical protein K8R - variola major virus  
C;Species: variola major virus  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T28500  
R;Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin  
Nature 366, 748-751, 1993  
A;Title: Potential virulence determinants in terminal regions of variola smallpox virus;  
A;Reference number: Z20488; MUID:94088747; PMID:8264798  
A;Accession: T28500  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-676 <MAS>  
A;Cross-references: UNIPROT:P33051; EMBL:L22579; NID:G623595; PIDN:AAA60810.1; PID:G43891  
A;Experimental source: strain Bangladesh-1975  
C;Superfamily: vaccinia virus I8 protein



Query Match 96.7%; Score 29; DB 2; Length 676;  
 Best Local Similarity 80.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:|  
 Db 632 DTYIH 636

RESULT 16  
 D72158  
 L8R protein - variola minor virus (strain Garcia-1966)  
 C:Species: variola minor virus  
 C>Date: 24-Nov-1999 #sequence\_revision 24-Nov-1999 #text\_change 17-Nov-2000  
 C:Accession: D72158  
 R;Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.P.; Massung, R.F.; Lopaz  
 submitted to GenBank, March 1998  
 A:Description: Analysis of the complete coding sequence of DNA of alastrim variola minor  
 A:Reference number: A72150  
 A:Accession: D72158  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-682 <SHC>  
 A:Cross-references: GB:Y16780; NID:G5830555; PIDN:CAB54662.1; PID:G5830623  
 A:Experimental source: strain Garcia-1966  
 C:Genetics:  
 A:Gene: L8R  
 C:Superfamily: vaccinia virus I8 protein

Query Match 96.7%; Score 29; DB 2; Length 682;  
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:|  
 Db 638 DTYIH 642

RESULT 17  
 H97827  
 organic solvent tolerance protein homolog [imported] - Rickettsia conorii (strain Malish  
 C:Species: Rickettsia conorii  
 C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
 C:Accession: H97827  
 R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro  
 Science 293, 2093-2098, 2001  
 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
 A:Reference number: A97700; MUID:21442074; PMID:11557893  
 A:Accession: H97827  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-713 <KUR>  
 A:Cross-references: UNIPROT:Q92GU8; GB:AE006914; PIDN:AAL03562.1; PID:G15620141; GSPDB:G  
 C:Genetics:  
 A:Gene: osta

Query Match 96.7%; Score 29; DB 2; Length 713;  
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:|  
 Db 161 DTYIH 165

RESULT 18  
 T50068  
 Probable ATP-dependent RNA helicase [imported] - fission yeast (Schizosaccharomyces pombe  
 C:Species: Schizosaccharomyces pombe  
 C>Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jul-2004  
 C:Accession: T50068  
 R;Seeger, K.; Harris, D.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, November 1999

A:Reference number: Z25033  
 A:Accession: T50068  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-735 <SEE>  
 A:Cross-references: UNIPROT:Q9UTP9; EMBL:AL132839; PIDN:CAB60250.1; GSPDB:GN00066; SPDB:  
 A:Experimental source: strain 972h(-); cosmid ci093  
 C:Genetics:  
 A:Gene: SPDB:SPAC1093.05  
 A:Map position: 1

Query Match 96.7%; Score 29; DB 2; Length 735;  
 Best Local Similarity 80.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:|  
 Db 367 DTYIH 371

RESULT 19  
 S56805  
 Probable RNA helicase CMA - Yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein J1250; protein YJL033w  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 08-Jul-1995 #sequence\_revision 08-Sep-1995 #text\_change 09-Jul-2004  
 C:Accession: S56805; C34848  
 R;Pohl, T.M.; Aljinovic, G.  
 submitted to the Protein Sequence Database, September 1995  
 A:Reference number: S56793  
 A:Accession: S56805  
 A:Molecule type: DNA  
 A:Residues: 1-770 <TOV>  
 A:Cross-references: UNIPROT:P20448; EMBL:Z49308; NID:G1008154; PIDN:CAA89324.1; PID:G100  
 R;Chang, T.H.; Arenas, J.; Abelson, J.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 1571-1575, 1990  
 A:Title: Identification of five putative yeast RNA helicase genes.  
 A:Reference number: A34848; MUID:90160368; PMID:2406722  
 A:Accession: C34848  
 A:Molecule type: DNA  
 A:Residues: 192-380, 'I', 382 <CHA>  
 C:Genetics:  
 A:Gene: SGD:HCA4  
 A:Cross-references: SGD:S0003570; MIPS:YJL033w  
 A:Map position: 10L  
 C:Keywords: ATP; nucleotide binding; P-loop  
 F;85-92/Region: nucleotide-binding motif A (P-loop)  
 F;190-195/Region: nucleotide-binding motif B  
 F;194-197/Region: DEAD motif

Query Match 96.7%; Score 29; DB 2; Length 770;  
 Best Local Similarity 80.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:|  
 Db 375 DTYIH 379

RESULT 20  
 T08851  
 Down syndrome cell adhesion protein 1 - human (fragment)  
 N:Alternate names: Down syndrome cell adhesion molecule  
 C:Species: Homo sapiens (man)  
 C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 11-Jun-1999  
 C:Accession: T08851  
 R;Yamakawa, K.; Huo, Y.K.; Haendel, M.A.; Hubert, R.; Chen, X.N.; Lyons, G.E.; Korenberg  
 submitted to the EMBL Data Library, September 1997  
 A:Description: DSCAM: A novel member of the immunoglobulin superfamily maps in a down sy  
 A:Reference number: Z16495  
 A:Accession: T08851  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA

A;Residues: 1-1896 <YAM>  
A;Cross-references: EMBL:AF023449; NID:g3169765; PID:g3169766  
A;Experimental source: brain; developmental stage: 14 weeks; fetal  
C;Genetics:  
A;Gene: DSCAM  
A;Map position: 21q22  
A;Note: derived from alternately-spliced mRNA  
C;Function:  
A;Description: involved in nervous system development  
C;Keywords: alternative splicing

Query Match 96.7%; Score 29; DB 2; Length 1896;  
Best Local Similarity 80.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||:|  
Db 506 DTYIH 510

RESULT 21  
T18497  
hypothetical protein C0780w - malaria parasite (Plasmodium falciparum)  
C;Species: Plasmodium falciparum  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T18497  
R;Lawson, D.; Bowman, S.; Barrell, B.  
submitted to the EMBL Data Library, August 1997  
A;Reference number: Z18935  
A;Accession: T18497  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-2657 <LAW>  
A;Cross-references: UNIPROT:O77380; EMBL:Z98551; NID:e1331903; PIDN:CAB11136.1  
C;Genetics:  
A;Map position: 3  
A;Note: C0780w

Query Match 96.7%; Score 29; DB 2; Length 2657;  
Best Local Similarity 80.0%; Pred. No. 5.5e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||:|  
Db 2027 DTYIH 2031

RESULT 22  
T18414  
protein g377 - malaria parasite (Plasmodium falciparum)  
C;Species: Plasmodium falciparum  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T18414  
R;Handman, E.; Oeborn, A.H.; Symons, F.; Van Driel, R.; Cappai, R.  
Mol. Biochem. Parasitol. 74, 143-156, 1995  
A;Title: The Leishmania promastigote surface antigen 2 complex is differentially expressed  
A;Reference number: Z18933; MUID:96360472; PMID:8719156  
A;Accession: T18414  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-3119 <HAN>  
A;Cross-references: UNIPROT:Q25857; EMBL:L04161; NID:g309687; PID:g309688; PIDN:AAC37257

Query Match 96.7%; Score 29; DB 2; Length 3119;  
Best Local Similarity 80.0%; Pred. No. 6.5e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||:|  
Db 2941 DTYIH 2945

RESULT 23

T08618  
intrinsic factor-B12 receptor CUBILIN precursor - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C;Accession: T08618  
R;Moestrup, S.K.; Kozyraki, R.; Kristiansen, M.; Kayesen, J.H.; Rasmussen, H.H.; Brault, J. Biol. Chem. 273, 5235-5242, 1998  
A;Title: The intrinsic factor-vitamin B12 receptor and target of teratogenic antibodies  
A;Reference number: Z16459; MUID:98148073; PMID:9478979  
A;Accession: T08618  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-3623 <MOE>  
A;Cross-references: UNIPROT:O70244; EMBL:AF022247; NID:g3834379; PIDN:AAC71661.1; PID:g3834379  
C;Genetics:  
A;Gene: CUBILIN  
C;Superfamily: intrinsic factor-B12 receptor cubilin; EGF homology  
C;Keywords: egg yolk; endocytosis; glycoprotein; intestine; kidney; peripheral membrane  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-3623/Product: intrinsic factor-B12 receptor CUBILIN #status predicted <MAT>  
F;133-164/Domain: EGF homology <EGF1>  
F;436-467/Domain: EGF homology <EGF>

Query Match 96.7%; Score 29; DB 2; Length 3623;  
Best Local Similarity 80.0%; Pred. No. 7.6e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||:|  
Db 492 DTYIH 496

RESULT 24  
S36377  
Ig heavy chain V region - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 23-Jul-1999  
C;Accession: S36377; S33392  
R;Ansell, K.H.  
submitted to the EMBL Data Library, April 1993  
A;Reference number: S36376  
A;Accession: S36377  
A;Molecule type: mRNA  
A;Residues: 1-51 <ANS>  
A;Cross-references: EMBL:X73023; NID:g295891; PIDN:CAA51506.1; PID:g939949  
R;Kettleborough, C.A.; Saldanha, J.; Ansell, K.H.; Bendig, M.M.  
Eur. J. Immunol. 23, 206-211, 1993  
A;Title: Optimization of primers for cloning libraries of mouse immunoglobulin genes using  
A;Reference number: S33391; MUID:93122092; PMID:8419173  
A;Accession: S33392  
A;Molecule type: mRNA  
A;Residues: 1-40, 'H' <KET>  
A;Cross-references: EMBL:X73023  
A;Experimental source: strain BALB/c  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotrimer; immunoglobulin

Query Match 90.0%; Score 27; DB 2; Length 51;  
Best Local Similarity 80.0%; Pred. No. 26;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
|||:|  
Db 31 DTYMH 35

RESULT 25  
H46200  
retrovirus-related reverse transcriptase homolog - ginkgo retrotransposon copia-like (fraxinus)  
C;Species: Ginkgo biloba (ginkgo)  
C;Date: 24-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C;Accession: H46200  
R;Voytas, D.F.; Cummings, M.P.; Konieczny, A.; Ausubel, F.M.; Rodermel, S.R.

Proc. Natl. Acad. Sci. U.S.A. 89, 7124-7128, 1992  
 A:Title: copia-like retrotransposons are ubiquitous among plants.  
 A:Reference number: A46200; MUID:92357784; PMID:1379734

A:Accession: H46200  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-88 <VO>  
 A:Cross-references: UNIPROT:Q06290; GB:M94475; NID:q439420; PIDN:AAA33351.1; PID:g168192  
 A:Note: sequence extracted from NCBI backbone (NCBIP:112113)  
 C:Superfamily: retrovirus-related polyprotein

Query Match 90.0%; Score 27; DB 2; Length 88;  
 Best Local Similarity 80.0%; Pred. No. 46;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:  
 Db 52 DTYLH 56

#### RESULT 26

A49715  
 Ig kappa chain V region (anti-glucuronoxylomannan antibody 439) - mouse  
 C:Species: Mus musculus (house mouse)

C:Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 21-Jan-2000  
 C:Accession: A49715

R:Otteson, E.W.; Welch, W.H.; Kozel, T.R.

J. Biol. Chem. 269, 1858-1864, 1994

A:Title: Protein-polysaccharide interactions. A monoclonal antibody specific for the cap  
 A:Reference number: A49715; MUID:94124530; PMID:8294434

A:Accession: A49715

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-112 <OTT>

A:Cross-references: GB:U01240; NID:g402265; PIDN:AAD10617.1; PID:g402266

A:Note: authors translated the codon CTT for residue 99 as Val

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:16-95/Domain: immunoglobulin homology <IMM>

Query Match 90.0%; Score 27; DB 2; Length 112;  
 Best Local Similarity 80.0%; Pred. No. 59;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:  
 Db 35 DTYLH 39

#### RESULT 27

A47271  
 nitrophenyl phosphonate-specific antibody 48G7 heavy chain VDJ - synthetic (fragment)

C:Species: synthetic

A:Note: Mus musculus (house mouse) gene engineered and expressed in Escherichia coli

C:Date: 21-Sep-1993 #sequence\_revision 11-Aug-1995 #text\_change 11-Aug-1995

C:Accession: A47271

R:Lesley, S.A.; Patten, P.A.; Schultz, P.G.

Proc. Natl. Acad. Sci. U.S.A. 90, 1160-1165, 1993

A:Title: A genetic approach to the generation of antibodies with enhanced catalytic acti

A:Reference number: A47271; MUID:93165660; PMID:8094556

A:Accession: A47271

A:Molecule type: DNA; protein

A:Residues: 1-114 <LRS>

A:Note: sequence extracted from NCBI backbone (NCBIN:124854, NCBIP:124855)

A:Note: parts of this sequence were determined by protein sequencing

F:22-96/Disulfide bonds: #status predicted

Query Match 90.0%; Score 27; DB 4; Length 114;  
 Best Local Similarity 80.0%; Pred. No. 60;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:  
 Db 30 DTYMH 34

#### RESULT 30

S17586

Ig heavy chain V region (E8) - mouse (fragment)

C:Species: Mus musculus (house mouse)

Db 31 DTYMH 35

#### RESULT 28

S03482

Ig heavy chain V-D-J region (hybridoma G8 Ad 3.8) - mouse (fragment)

C:Species: Mus musculus (house mouse)

A:Variety: strain BALB/c

C:Date: 26-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 21-Jan-2000

C:Accession: S03482; S07453

R:Rocca-Serra, J.; Matthes, H.W.; Kaartinen, M.; Milstein, C.; Theze, J.; Fougereau, M.

EMBO J. 2, 867-872, 1983

A:Title: Analysis of antibody diversity: V-D-J mRNA nucleotide sequence of four anti-GAT

hypervariable regions

A:Reference number: S03471; MUID:84057768; PMID:6416834

A:Accession: S03482

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 10-115 <ROC1>

A:Cross-references: EMBL:X03219

A:Note: this sequence was determined from the differentiated gene

R:Rocca-Serra, J.; Mazie, J.C.; Moinier, D.; Leclercq, L.; Somme, G.; Theze, J.; Fougere

J. Immunol. 129, 2554-2558, 1982

A:Title: The limited diversity of the mouse gamma-chains anti-GAT repertoire does not se

A:Reference number: S07453; MUID:83058021; PMID:6815271

A:Accession: S07453

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-43 <ROC2>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 90.0%; Score 27; DB 2; Length 115;  
 Best Local Similarity 80.0%; Pred. No. 60;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:  
 Db 31 DTYMH 35

#### RESULT 29

S24289

Ig gamma chain V region (JS34/32) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Jun-2000

C:Accession: S24289

R:Monchamont, B.

submitted to the EMBL Data Library, September 1991

A:Description: Cloning and sequencing of the cDNA coding for the variable regions of the

A:Reference number: S24287

A:Accession: S24289

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-116 <MON>

A:Cross-references: EMBL:X62705; NID:g51690; PIDN:CAA44584.1; PID:g13333963

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: immunoglobulin

F:14-97/Domain: immunoglobulin homology <IMM>

Query Match 90.0%; Score 27; DB 2; Length 116;  
 Best Local Similarity 80.0%; Pred. No. 61;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
 |||:  
 Db 30 DTYMH 34

C;Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 21-Jan-2000  
A;Accession: S17586  
R;Mylavanam, S.E.; Paterson, Y.; Kaiser, K.; Bowdish, K.; Getzoff, E.D.  
J. Mol. Biol. 221, 455-462, 1991  
A;Title: Biochemical implications from the variable gene sequences of an anti-cytochrome forms.  
A;Reference number: S17586; MUID:92015240; PMID:1656053  
A;Accession: S17586  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-117 <MVL>  
A;Cross-references: EMBL:X60683; NID:g51820; PIDN:CAA43095.1; PID:g51821  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 90.0%; Score 27; DB 2; Length 117;  
Best Local Similarity 80.0%; Pred. No. 61;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 31 DTYMH 35

RESULT 31  
S03484  
Ig heavy chain V-D-J region (hybridoma G7 Ab 2.9) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
A;Variety: strain BALB/c  
C;Date: 26-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 21-Jan-2000  
A;Accession: S03484; S07453  
R;Rocca-Serra, J.; Matthes, H.W.; Kaartinen, M.; Milstein, C.; Theze, J.; Fougereau, M.  
EMBO J. 2, 867-872, 1983  
A;Title: Analysis of antibody diversity: V-D-J mRNA nucleotide sequence of four anti-GAT hypervariable regions.  
A;Reference number: S03471; MUID:84057768; PMID:6416834  
A;Accession: S03484  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 10-120 <ROC1>  
A;Cross-references: EMBL:X07144  
A;Note: this sequence was determined from the differentiated gene  
R;Rocca-Serra, J.; Mazie, J.C.; Moinier, D.; Leclercq, L.; Somme, G.; Theze, J.; Fougere J. Immunol. 129, 2554-2558, 1982  
A;Title: The limited diversity of the mouse gamma-chains anti-GAT repertoire does not se  
A;Reference number: S07453; MUID:83058021; PMID:6815271  
A;Accession: S07453  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-43 <ROC2>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 90.0%; Score 27; DB 2; Length 120;  
Best Local Similarity 80.0%; Pred. No. 63;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 31 DTYMH 35

RESULT 32  
S03471  
Ig heavy chain V-D-J region (hybridoma G5 Bb 2.2) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Jun-2000  
A;Accession: S03471; S07453  
R;Rocca-Serra, J.; Matthes, H.W.; Kaartinen, M.; Milstein, C.; Theze, J.; Fougereau, M.  
EMBO J. 2, 867-872, 1983  
A;Title: Analysis of antibody diversity: V-D-J mRNA nucleotide sequence of four anti-GAT hypervariable regions.

Query Match 90.0%; Score 27; DB 2; Length 120;  
Best Local Similarity 80.0%; Pred. No. 63;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 31 DTYMH 35

RESULT 33  
S06823  
Ig heavy chain V region (clone I1c) - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 21-Jan-2000  
A;Accession: S06823  
R;Miller III, A.; Glasel, J.A.  
J. Mol. Biol. 209, 763-778, 1989  
A;Title: Comparative sequence and immunochemical analyses of murine monoclonal anti-morpi  
A;Reference number: S06815; MUID:90064531; PMID:2555519  
A;Accession: S06823  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-122 <MIL>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-99/Domain: immunoglobulin homology <IMM>

Query Match 90.0%; Score 27; DB 2; Length 122;  
Best Local Similarity 80.0%; Pred. No. 64;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 31 DTYMH 35

RESULT 34  
PH1403  
Ig heavy chain V region (clone micro m- 46-6, 46-12) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Jan-2000  
A;Accession: PH1403  
R;Shirasawa, T.; Miyazoe, I.; Hagiwara, S.; Kimoto, H.; Shigemoto, K.; Taniguchi, M.; Tai J. Exp. Med. 176, 1209-1214, 1992  
A;Title: Heavy chain variable (VH) region diversity generated by VH gene replacement in ia virus.  
A;Reference number: PH1403; MUID:93018837; PMID:1402663  
A;Accession: PH1403  
A;Molecule type: DNA  
A;Residues: 1-123 <SHI>  
C;Genetics:  
A;Introns: 16/1  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;35-118/Domain: immunoglobulin homology <IMM>

Query Match 90.0%; Score 27; DB 2; Length 123;  
Best Local Similarity 80.0%; Pred. No. 65;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 51 DTYMH 55

RESULT 35  
S52449  
Ig kappa chain V region - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 08-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 21-Jan-2000  
C;Accession: S52449  
R;Bardos, J.; Kraehenbuhl, J.P.  
submitted to the EMBL Data Library, November 1994  
A;Description: Specific amplification by the polymerase chain reaction of rearranged germline DNA  
A;Accession: S52449  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-131 <BER>  
A;Cross-references: EMBL:X82691; NID:G673446; PIDN:CAA58012.1; PID:G673447  
C;Genetics:  
A;Introns: 16/1  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;35-114/Domain: immunoglobulin homology <IMM>

Query Match 90.0%; Score 27; DB 2; Length 131;  
Best Local Similarity 80.0%; Pred. No. 69;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 54 DTYLH 58

RESULT 36  
AG1191  
B. subtilis CspR protein, rRNA methylase homolog lmo0935 [imported] - Listeria m  
C;Species: Listeria monocytogenes  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AG1191  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AG1191  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-169 <GLA>  
A;Cross-references: UNIPROT:Q8Y8G9; GB:NC\_003210; PIDN:CAC99013.1; PID:gl6410337; GSPDB: A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: lmo0935  
C;Superfamily: Chlamydomydia pneumoniae rRNA methylase

Query Match 90.0%; Score 27; DB 2; Length 169;  
Best Local Similarity 80.0%; Pred. No. 90;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 27 DTYLH 31

RESULT 37

S29594  
Ig gamma chain (MM65) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999  
C;Accession: S29594  
R;Seymour, R.  
submitted to the EMBL Data Library, February 1991  
A;Reference number: S29593  
A;Accession: S29594  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-178 <SEY>  
A;Cross-references: EMBL:X57857; NID:G52590; PIDN:CAA40992.1; PID:G52591  
C;Keywords: immunoglobulin

Query Match 90.0%; Score 27; DB 2; Length 178;  
Best Local Similarity 80.0%; Pred. No. 94;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 44 DTYMH 48

RESULT 38  
S49220  
Ig gamma-1 chain - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 19-Mar-1997 #sequence\_revision 01-Aug-1997 #text\_change 21-Jan-2000  
C;Accession: S49220  
R;Kipp, B.; Becker, W.P.; Schlaak, M.M.  
submitted to the EMBL Data Library, September 1994  
A;Description: Cloning and expression of a recombinant mouse Fab-fragment recognizing a  
A;Reference number: S49220  
A;Accession: S49220  
A;Molecule type: mRNA  
A;Residues: 1-221 <KIP>  
A;Cross-references: EMBL:Z37502; NID:G541778; PIDN:CAA85732.1; PID:G541779  
A;Experimental source: strain Balb/c  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: immunoglobulin  
F;1-120/Domain: V region #status predicted <VRG>  
F;121-221/Domain: C region #status predicted <CRG>  
F;139-203/Domain: immunoglobulin homology <IMM>

Query Match 90.0%; Score 27; DB 2; Length 221;  
Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5  
Db 31 DTYMH 35

RESULT 39  
A56446  
Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/c)  
C;Species: Mus musculus (house mouse)  
C;Date: 19-Jan-1996 #sequence\_revision 19-Jan-1996 #text\_change 16-Aug-1996  
C;Accession: A56446  
R;Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.  
J. Biol. Chem. 270, 7829-7835, 1995  
A;Title: A high affinity digoxin-binding protein displayed on M13 is functionally identical  
A;Reference number: A56446; MUID:95229583; PMID:7713873  
A;Accession: A56446  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-268 <THN>  
A;Cross-references: GB:U20617  
C;Keywords: heterotetramer; immunoglobulin

Query Match 90.0%; Score 27; DB 2; Length 268;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5

|||:|

Db 33 DTYVH 37

RESULT 40

S72886

hypothetical protein B2168\_F3\_l30 - Mycobacterium leprae

C;Species: Mycobacterium leprae

C;Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004

C;Accession: S72886

R;Smith, D.R.; Robison, K.

submitted to the EMBL Data Library, November 1993

A;Description: Mycobacterium leprae cosmid B2168.

A;Reference number: S72586

A;Accession: S72886

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-308 <SMI>

A;Cross-references: UNIPROT:Q49807; EMBL:U00018; NID:G467037; PIDN:AAAL7222.1; PID:G4670

Query Match

Best Local Similarity 90.0%; Score 27; DB 2; Length 308;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5

|||:|

Db 281 DTYVH 285

Search completed: June 3, 2005, 16:02:32

Job time : 10.7879 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 15:40:50 ; Search time 42.4242 Seconds  
(without alignments)  
60.352 Million cell updates/sec

Title: US-10-089-452-21

Perfect score: 30

Sequence: 1 DTVVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	127	2	Q8CM07 streptococ
2	30	100.0	141	2	Q8BGR7 m mus muscu
3	30	100.0	142	2	Q919P2 culex nigri
4	30	100.0	195	2	Q8DXU6 Q8dxu6 streptococ
5	30	100.0	195	2	Q8E3G5 streptococ
6	30	100.0	195	2	Q8PED9 mus musculu
7	30	100.0	206	2	Q665V0 Q665v0 yersinia ps
8	30	100.0	207	2	Q8Z1G3 Q8zi63 yersinia pe
9	30	100.0	220	2	Q8F821 Q8p821 xenopus tro
10	30	100.0	223	2	Q8KRF3 Q8krf3 arthrobacte
11	30	100.0	257	2	Q969T0 Q969t0 homo sapien
12	30	100.0	300	2	Q8EFH0 Q8efh0 shewanella
13	30	100.0	304	2	Q658P1 Q658p1 homo sapien
14	30	100.0	304	2	Q8NXY2 Q8nxy2 mus musculu
15	30	100.0	305	2	Q8C4Q6 Q8c4q6 mus musculu
16	30	100.0	305	2	Q99K80 Q99k80 mus musculu
17	30	100.0	305	2	Q6PAW0 Q6paw0 xenopus lae
18	30	100.0	305	2	Q6PB19 Q6pb19 xenopus lae
19	30	100.0	306	2	Q96BJ3 Q96bj3 homo sapien
20	30	100.0	306	2	Q9H9E8 Q9h9e8 homo sapien
21	30	100.0	325	2	P87255 P87255 aspergillus
22	30	100.0	342	2	Q937G0 Q937g0 lactobacill
23	30	100.0	345	2	Q7L3J6 Q7l3j6 homo sapien
24	30	100.0	345	2	Q810R6 Q810r6 mus musculu
25	30	100.0	373	2	Q69849 Q69849 streptomyc
26	30	100.0	377	2	Q8C1Y2 Q8c1y2 mus musculu
27	30	100.0	395	2	Q9ZWW7 Q9zww7 corynebacte
28	30	100.0	398	2	Q8EMW7 Q8emw7 oceanobacil
29	30	100.0	438	2	Q6LQJ2 Q6lqj2 photobacter
30	30	100.0	446	2	Q95QJ6 Q95qj6 caenorhabdi
31	30	100.0	456	2	Q8TA83 Q8ta83 caenorhabdi

32	30	100.0	492	2	Q6NIT6 Q6nitr6 corynebacte
33	30	100.0	497	2	Q8XRR2 Q8xrr2 ralstonia s
34	30	100.0	543	2	Q935V4 Q935v4 novosphingo
35	30	100.0	552	2	Q71JB5 Q71jb5 homo sapien
36	30	100.0	573	2	Q9GV07 Q9gv07 dugesia dor
37	30	100.0	630	2	Q9Y7A2 Q9y7a2 coprinus ci
38	30	100.0	673	2	Q70EW1 Q70ew1 glycine max
39	30	100.0	674	1	AMO.PEA Q43077 ptaum sativ
40	30	100.0	695	1	EPGI_SYNY3 P28371 synechocyst
41	30	100.0	729	2	Q6C1Z8 Q6c1z8 kluyveromyc
42	30	100.0	738	2	Q7YVG6 Q7yvg6 cryptospori
43	30	100.0	759	2	Q86V87 Q86v87 homo sapien
44	30	100.0	822	2	Q62287 Q62287 mus musculu
45	30	100.0	943	1	YLM5_CABEL P34408 caenorhabdi

## ALIGNMENTS

### RESULT 1

Q8CM07 Q8CM07 PRELIMINARY; PRT; 127 AA.  
AC Q8CM07;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Hypothetical protein gbs0408 (Hypothetical protein gbs0738)  
DE (Hypothetical protein gbs0971).  
GN OrderedLocNames=gbs0408, gbs0738, gbs0971;  
OS Streptococcus agalactiae (serotype III).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=216495;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NEM316 / Serotype III;  
RX MEDLINE=22242508; PubMed=12354221;  
RA Glaser P., Rusniok C., Buchrieser C., Chevalier P., Frangeul L.,  
RA Meadek T., Zouine M., Couve E., Lalloui L., Poyart C., Trieu-Cuot P.,  
RA Kunst F.;  
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing  
RT invasive neonatal disease."  
RL Mol. Microbiol. 45:1499-1513 (2002).  
DR EMBL; AL766845; CAD46052.1; -  
DR EMBL; AL766846; CAD46382.1; -  
DR EMBL; AL766848; CAD46630.1; -  
DR SagalList; gbs0408; -  
DR SagalList; gbs0738; -  
DR SagalList; gbs0971; -  
DR InterPro; IPR010724; RepA\_N.  
DR InterPro; IPR009058; Wing\_hlx\_DNA\_bnd.  
DR Pfam; PF06970; RepA\_N; 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 127 AA; 14822 MW; E1BD5489D2888B88 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 127;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5

Db 68 DTVVH 72

### RESULT 2

Q8BGR7 Q8BGR7 PRELIMINARY; PRT; 141 AA.  
AC Q8BGR7;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Mus musculus 13 days embryo lung cDNA, RIKEN full-length enriched  
DE library, clone:D430028H19 product:hypothetical protein, full insert

sequence (Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched library, clone:110051C19 product:hypothetical protein, full insert sequence.) (Fragment).

GN Name=2610208M17Rik;

OS Mus musculus (Mouse);

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI\_TaxID=10090;

[1]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Lung, and Whole body;

RA MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9; Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44(1999).

[2]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Lung, and Whole body;

RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500; Carninci P., Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

[3]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Lung, and Whole body;

RA The FANTOM Consortium;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

[4]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Lung, and Whole body;

RA MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100; Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).

[5]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Lung, and Whole body;

RA MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600; Shibata K., Itoh M., Aizawa K., Nagao K., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).

[6]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Lung;

RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayaehida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Nishihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

[7]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,

RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK085039; BAC39347.1; -

DR EMBL; AK027971; BAC25682.1; -

DR MGD; MGI:1919737; 2610208M17Rik.

KW Hypothetical protein.

SQ SEQUENCE 141 AA; 16458 MW; 79340F6C9267CCEP CRC64;

Query Match 100.0%; Score 30; DB 2; Length 141;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVVH 5

Db 50 DTVVH 54

RESULT 3

Q919P2 PRELIMINARY; PRT; 142 AA.

AC Q919P2; 2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Hypothetical protein CUN027.

GN Name=CUN027;

OS Culex nigripalpus baculovirus.

OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae.

OX NCBI\_TaxID=130556;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21488685; PubMed=11602755;

RX DOI=10.1128/JVI.75.22.11157-11165.2001;

RA Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A., Becnel J.J., Rock D.L., Kutish G.F.;

RT "Genome sequence of a baculovirus pathogenic for Culex nigripalpus.";

RL J. Virol. 75:11157-11165(2001).

DR EMBL; AF403738; AAK94105.1; -

KW Hypothetical protein.

SQ SEQUENCE 142 AA; 16456 MW; 80E2540E32E519D2 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 142;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVVH 5

Db 36 DTVVH 40

RESULT 4

Q8DXU6 PRELIMINARY; PRT; 195 AA.

AC Q8DXU6; 2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Exonuclease.

GN OrderedLocutusNames=SAG1750;

OS Streptococcus agalactiae (serotype V).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=216466;

RN [1]

RP SEQUENCE FROM N.A.



```

RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547; DOI=10.1073/pnas.182380799;
RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.N.,
RA Wessels L.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
RA Carly H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL; AB014271; AAN00613.1; -.
DR TIGR; SAG1750; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR InterPro; IPR006055; Exonuclease.
DR Pfam; PF00929; Exonuc_X-T; 1.
DR SMART; SM00479; EXOIII; 1.
DR Complete proteome.
SQ SEQUENCE 195 AA; 21939 MW; DB3ABF934B668819 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 72; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0;

Qy 1 DTYVH 5
Db 41 DTYVH 45

RESULT 5
Q8E3G5 PRELIMINARY; PRT; 195 AA.
AC Q8E3G5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein gbs1794.
GN OrderedLocusNames=gbs1794;
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA Meadek T., Zouine M., Couve E., Lalloui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL; AL76853; CAD47453.1; -.
DR Sgallist; gbs1794; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR InterPro; IPR006055; Exonuclease.
DR Pfam; PF00929; Exonuc_X-T; 1.
DR SMART; SM00479; EXOIII; 1.
DR Complete proteome.
SQ SEQUENCE 195 AA; 21851 MW; 8B826727072409A1 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 72; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0;

Qy 1 DTYVH 5
Db 41 DTYVH 45

```

```

RESULT 6
Q6PED9 PRELIMINARY; PRT; 195 AA.
AC Q6PED9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smillie D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marz M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058108; AAH58108.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 195 AA; 22227 MW; 233FD0CB8EF9E9FA CRC64;

Query Match 100.0%; Score 30; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 72; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0;

Qy 1 DTYVH 5
Db 12 DTYVH 16

```

```

RESULT 7
Q665V0 PRELIMINARY; PRT; 206 AA.
AC Q665V0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Conserved hypothetical membrane protein. precursor.
GN ORFNames=YPTB3410;
OS Yersinia pseudotuberculosis IP 32953.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=273123;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP 32953;

```

```

RX PubMed=15358858;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin L.V.,
RA Brubaker R.R., Fowler J., Hinnebusch B.J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francique V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the genome evolution of Yersinia pestis through whole
RT genome comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH22648.1; -.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR003646; SH3_bac.
DR SMART; SM00287; SH3b; 1.
KW Hypothetical protein; Signal.
FT SIGNAL 1 23 Potential.
SQ SEQUENCE 206 AA; 23128 MW; E5BFD0D1A6PB4504 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 33 DTYVH 37
|||||

RESULT 8
ID Q82I63 PRELIMINARY; PRT; 207 AA.
AC Q82I63; Q74RQ6; Q7CGG7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative exported protein (Hypothetical protein y3528).
GN OrderedLocusNames=YP2966, YP00651, y3528;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Tittball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyscon P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ414144; CAC89505.1; -.

```

```

DR EMBL; AE013956; AAM87076.1; -.
DR EMBL; AE017138; AAS63145.1; -.
DR PTR; AF0080; AF0080.
DR SMART; SM00287; SH3b; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 207 AA; 23217 MW; 740EB839DF5DFCA5 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 33 DTYVH 37
|||||

RESULT 9
ID Q6P821 PRELIMINARY; PRT; 220 AA.
AC Q6P821;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC76000.
GN Name=MGC76000;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061410; AH461410.1; -.
KW Hypothetical protein.
SQ SEQUENCE 220 AA; 25093 MW; 296A99D83C858269 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 220;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 129 DTYVH 133
|||||

RESULT 10

```

```

O8KRF3
ID Q8KRF3 PRELIMINARY; PRT; 223 AA.
AC Q8KRF3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transcriptional regulator.
DE Name=dbiAC22;
GN Arthrobacter sp. C2-2.
OS Arthrobacter; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=192168;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2-2;
RA Karasova P., Strnad H., Spiwok V., Mala S., Kralova B., Russell N.J.;
RT "The cloning, purification and characterization of a cold active beta-
galactosidase from the psychrotolerant Antarctic bacterium
RT Arthrobacter sp. C2-2";
RL Enzyme Microb. Technol. 33:836-844 (2003).
CC -1- SIMILARITY: Contains 1 response regulatory domain.
DR EMBL; AJ457162; CAD29778.1; -.
DR HSSP; Q9WYN0; 1KGS.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR009059; bi_resp_registr_C.
DR InterPro; IPR011006; CheY-like.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001867; Trans_reg_C.
DR Pfam; PF00072; Response_reg; 1.
DR Pfam; PF00486; Trans_reg_C; 1.
DR ProDom; PD000039; Response_reg; 1.
DR ProDom; PD000329; Trans_reg_C; 1.
DR SMART; SM00448; REC_1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
KW DNA-binding; Phosphorylation; Sensory transduction; Transcription;
KW Transcription regulation.
SQ SEQUENCE 223 AA; 24245 MW; 5647ABE5F693E080 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 83; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0;

Qy 1 DTYVH 5
Db 194 DTYVH 198

RESULT 11
Q969T0 ID Q969T0 PRELIMINARY; PRT; 257 AA.
AC Q969T0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE RAI16 protein (Fragment).
GN Name=RAI16;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

Query Match 100.0%; Score 30; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 74 DTYVH 78

RESULT 12
Q8EFH0 ID Q8EFH0 PRELIMINARY; PRT; 300 AA.
AC Q8EFH0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein SO2002.
GN OrderedLocuNames=SO2002;
OS Shewanella oneidensis
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Shewanellaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RA MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA Clayton R.A., Meyer T., Tsapin A., Scott J., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Unayam L.A.,
RA White O., Wolf A.M., Vamathevan J.J., Weidman J.P., Imbraim M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Uterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
Shewanella oneidensis";
RL Nat. Biotechnol. 20:1118-1123 (2002).
DR EMBL; AE015642; AAN55053.1; -.
DR TIGR; SO2002; -.
DR InterPro; IPR005639; endotoxin_N.
KW Complete proteome.
SQ SEQUENCE 300 AA; 33112 MW; 6AA0F8D62780E41E CRC64;

Query Match 100.0%; Score 30; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 74 DTYVH 78

```

```
Db          96 DTYVH 100

RESULT 13
Q658P1      PRELIMINARY;      PRT;      304 AA.
AC Q658P1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp666C185 (Fragment).
GN Name=DKFZp666C185;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stomach;
RG The German cDNA Consortium;
RA Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL833718; CAH56257.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 304 AA; 34805 MW; 6FC4889221B86E46 CRC64;

Query Match      100.0%; Score 30; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 213 DTYVH 217

RESULT 14
Q6NXY2      PRELIMINARY;      PRT;      304 AA.
AC Q6NXY2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE 2610208M17Rik protein (Fragment).
GN Name=2610208M17Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic Germ Cell;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
```

```
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic Germ Cell;
RA Strausberg R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC066829; AAH66829.1; -.
FT NON TER 1
SQ SEQUENCE 304 AA; 34799 MW; E3A328C478A7C024 CRC64;

Query Match      100.0%; Score 30; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 213 DTYVH 217

RESULT 15
Q8C4Q6      PRELIMINARY;      PRT;      305 AA.
AC Q8C4Q6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
DE library, clone:Cl30021P14 product:hypothetical protein, full insert
DE sequence.
GN Name=2610208M17Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakeguchi S., Ikegami T., Kashiwagi K.,
```

RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Orawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.,  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RL sequencing pipeline with 384 multicapillary sequencer.";   
 RN Genome Res. 10:1757-1771(2000).  
 [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu M., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Saeki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK081494; BAC38234.1; -  
 KW MGD; MGI:1919737; 2610208M17Rik.  
 SQ SEQUENCE 305 AA; 34888 MW; A285C564AF533D0 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVYH 5  
 Db 214 DTVYH 218

RESULT 16  
 Q99K80 PRELIMINARY; PRT; 305 AA.  
 AC Q99K80;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE RIKEN cDNA 2610208M17.  
 GN Name=2610208M17Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mammary tumor;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Small D.B., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mammary tumor;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Small D.B., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mammary tumor;  
 RA Strausberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC004835; AA04835.1; -  
 DR MGD; MGI:1919737; 2610208M17Rik.  
 SQ SEQUENCE 305 AA; 34860 MW; A285DA6924C11ED CRC64;

Query Match 100.0%; Score 30; DB 2; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVYH 5  
 Db 214 DTVYH 218

RESULT 17  
 Q6PAW0 PRELIMINARY; PRT; 305 AA.  
 AC Q6PAW0;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE MGC68703 protein.  
 GN Name=MGC68703;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Small D.B., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Small D.B., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Small D.B., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]

Query Match 100.0%; Score 30; DB 2; Length 305;

```
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5
Db 214 DTVVH 218

RESULT 18
Q6PB19 PRELIMINARY; PRT; 305 AA.
AC Q6PB19
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC68484 protein.
GN Name=MGC68484;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., McEwan P.J., McKernan K.J., Abramson R.D., Mullany S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC015535; AAH15535.1; -.
DR EMBL; BC067805; AAH67805.1; -.
KW Hypothetical protein.
SQ SEQUENCE 306 AA; 35023 MW; 48E563CEC398ECC CRC64;

Query Match 100.0%; Score 30; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5
Db 215 DTVVH 219

RESULT 20
Q9H9E8 PRELIMINARY; PRT; 306 AA.
AC Q9H9E8
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ12806.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

Query Match 100.0%; Score 30; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTVVH 5
Db 215 DTVVH 219

RESULT 20
Q9H9E8 PRELIMINARY; PRT; 306 AA.
AC Q9H9E8
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ12806.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```

RP SEQUENCE FROM N.A.
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosori T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kinata M., Watanabe M., Hiraoa K., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Mueshino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugeno J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK022868; BAB14281.1; -.
SQ SEQUENCE 306 AA; 34951 MW; B368D499B8E16D41 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 215 DTYVH 219
|||||

RESULT 21
P87255 ID P87255 PRELIMINARY; PRT; 325 AA.
AC P87255;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Orf1 3' of Ant1 5' insertion site protein.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=Ant1;
RX MEDLINE=96133683; PubMed=8552048;
RA Glayzer D.C., Roberts I.N., Archer D.B., Oliver R.P.;
RT "The isolation of Ant1, a transposable element from Aspergillus
RT niger.";
RL Mol. Gen. Genet. 249:432-438(1995).
DR EMBL; S80872; AAB50684.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR Pfam; PF00665; rve; 1.
SQ SEQUENCE 325 AA; 37250 MW; 63213AB354112C15 CRC64;

```

```

Query Match 100.0%; Score 30; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 131 DTYVH 135
|||||

RESULT 22
Q937G0 ID Q937G0 PRELIMINARY; PRT; 342 AA.
AC Q937G0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative transposase.
GN Name=ISLd13;
OS Lactobacillus delbrueckii (subsp. lactis).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=29397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15808;
RX MEDLINE=22635853; PubMed=12749837; DOI=10.1016/S0147-619X(03)00018-0;
RA Ravin V.K., Alatosava T.;
RT "Three new insertion sequence elements ISLd12, ISLd13, and ISLd14 in
RT Lactobacillus delbrueckii: isolation, molecular characterization, and
RT potential use for strain identification.";
RL Plasmid 49:253-268(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15808;
RX MEDLINE=21997078; PubMed=12002399;
RA Ravin V., Alatosava T.;
RT "A new insertion sequence element, ISLd11, in Lactobacillus
RT delbrueckii subsp. lactis ATCC 15808.";
RL Microbiol. Res. 157:109-114(2002).
DR EMBL; AJ316615; CAC79142.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
SQ SEQUENCE 342 AA; 39745 MW; 1B4B6A3FB99743C CRC64;

Query Match 100.0%; Score 30; DB 2; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 116 DTYVH 120
|||||

RESULT 23
Q7L3J6 ID Q7L3J6 PRELIMINARY; PRT; 345 AA.
AC Q7L3J6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE RAI16 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

```

```
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pehey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strauberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049609; AAH49609.1; -.
DR MGI; MGI:102518; Adam3.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR InterPro; IPR006585; ADAM_cysteine.
DR InterPro; IPR007042; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PSS0026; EGF_3; 1.
FT NON TER 1
SQ SEQUENCE 345 AA; 38958 MW; 98C14FE2206E0CA6 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 162 DTYVH 166

RESULT 24
Q810R6 PRELIMINARY; PRT; 345 AA.
AC Q810R6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Adam3 protein (Fragment).
GN Names=Adam3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pehey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
```

```
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strauberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049609; AAH49609.1; -.
DR MGI; MGI:102518; Adam3.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR InterPro; IPR006585; ADAM_cysteine.
DR InterPro; IPR007042; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PSS0026; EGF_3; 1.
FT NON TER 1
SQ SEQUENCE 345 AA; 38690 MW; B12E3C4947F23350 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 345;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 114 DTYVH 118

RESULT 25
O69849 PRELIMINARY; PRT; 373 AA.
AC O69849;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SCO6019.
GN ORFNames=SC13.07;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RC MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphry L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL Nature 417:141-147(2002).
DR EMBL; AL339126; CAAL9231.1; -.
DR PIR; T34699; T34699.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 373 AA; 39303 MW; 80405EB84CB7216E CRC64;

Query Match 100.0%; Score 30; DB 2; Length 373;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 89 DTYVH 93

RESULT 26
```



O8CLY2 PRELIMINARY; PRT; 377 AA.  
AC O8CLY2;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Mus musculus 1.5 years female mammary gland CRL-2116 JC cDNA, RIKEN  
DE full-length enriched library, clone:G43067P06 product:hypothetical  
DE ATP/GTP-binding site motif A (P-loop) containing protein, full insert  
DE sequence.  
GN Name=Raill6;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/cRos; TISSUE=Mammary gland;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/cRos; TISSUE=Mammary gland;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN PANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/cRos; TISSUE=Mammary gland;  
RA The PANTOM Consortium;  
RT the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/cRos; TISSUE=Mammary gland;  
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/cRos; TISSUE=Mammary gland;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ogawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/cRos; TISSUE=Mammary gland;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara K., Matsuura T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito K., Saitoh H., Sakai K., Sakazume N., Sano H.,  
RA Sagaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,

Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK090035; BAC41057.1; -;  
DR MGD; MGI:3036290; Rail6.  
KW Hypothetical protein.  
SQ SEQUENCE 377 AA; 41937 MW; A9DAA08B08A2C090 CRC64;  
Query Match 100.0%; Score 30; DB 2; Length 377;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DTYVH 5  
Db 270 DTYVH 274  
|||||  
RESULT 27  
Q92WV7 PRELIMINARY; PRT; 395 AA.  
AC Q92WV7;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Integrase.  
GN Name=int;  
OS Corynebacterium 304L.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
OX NCBI\_TaxID=86386;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99160765; PubMed=10049830; DOI=10.1006/viro.1998.9524;  
RA Moreau S., Le Marrec C., Blanco C., Trautwetter A.;  
RT "Analysis of the integration function of phi304L: an integrase module  
RT among corynebacteriophages.";  
RL Virology 255:150-159(1999).  
DR EMBL; Y18058; CAB38562.1; -;  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0015074; P:DNA integration; IEA.  
DR GO; GO:0006310; P:DNA recombination; IEA.  
DR InterPro; IPR011010; DNA\_brk\_join enz.  
DR InterPro; IPR002104; Phage integrase.  
DR Pfam; PF00589; Phage integrase; 1.  
SQ SEQUENCE 395 AA; 43592 MW; 173670D24F9EC89F CRC64;  
Query Match 100.0%; Score 30; DB 2; Length 395;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 DTYVH 5  
Db 355 DTYVH 359  
|||||  
RESULT 28  
Q8EMA7 PRELIMINARY; PRT; 398 AA.  
AC Q8EMA7;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Xaa-Pro aminopeptidase.  
GN OrderedLocustNames=OB2950;  
OS Oceanobacillus thelyensis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.  
OX NCBI\_TaxID=182710;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HTE831;  
RX MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkf526;  
RA Takami H., Takaki Y., Uchiyama I.;  
RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya  
RT Ridge and its unexpected adaptive capabilities to extreme  
RT environments.";

```

RL Nucleic Acids Res. 30:3927-3935 (2002).
DR EMBL; AP004603; BAC14906.1; -.
DR GO; GO:0008235; F:metallopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000994; Peptidase M24.
DR Pfam; PF00557; Peptidase_M24; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 398 AA; 44533 MW; 4BB261C898C33AFE CRC64;

Query Match 100.0%; Score 30; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 89 DTYVH 93

RESULT 29
Q6LQU2 Q6LQU2 PRELIMINARY; PRT; 438 AA.
AC Q6LQU2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative guanine aminohydrolase.
GN Name=SMB21293; OrderedLocusNames=PBPR1930;
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP SEQUENCE FROM N.A.
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
RA Cestaro A., Malacrida G., Simionati B., Cannata N., Bartlett D.,
RA Valle G.;
RT "Genome analysis of Photobacterium profundum reveals the complexity of
RT high pressure adaptations.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR378669; CAG20334.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR011550; Amidohydro_like.
DR InterPro; IPR011059; Metallo_hydrolase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR ProDom; PD001248; Amidohydro_like; 1.
KW Complete proteome.
SQ SEQUENCE 438 AA; 49941 MW; BA2408FBDAC77295 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 438;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 230 DTYVH 234

RESULT 30
Q95QJ6 Q95QJ6 PRELIMINARY; PRT; 446 AA.
AC Q95QJ6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 26, Last annotation update)
DE Dnaj domain (Prokaryotic heat shock protein) protein 10, isoform
DE b.
GN Name=dnj-10; ORFNames=F22B7.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peleoriniae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Anderson K.;
RT "The sequence of C. elegans cosmid F22B7.";
RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; L12018; AAK68343.1; -.
DR HSP; P08622; IXL.
DR WormBase; WBGene00001028; F22B7.5.
DR WormPep; F22B7.5b; CE27991.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR GO; GO:0006986; P:response to unfolded protein; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR002939; Dnaj_C.
DR InterPro; IPR001305; Dnaj_CXXCXGKG.
DR InterPro; IPR001623; Dnaj_N.
DR InterPro; IPR008971; HSP40_Dnaj_pep.
DR InterPro; IPR003095; Hsp_Dnaj.
DR Pfam; PF00226; Dnaj; 1.
DR Pfam; PF01556; Dnaj_C; 1.
DR PRINTS; PR00625; DNAPROTEIN.
DR SMART; SM00271; Dnaj; 1.
DR SMART; SM00343; Znf_C2HC; 2.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00636; DNAJ_1; 1.
DR PROSITE; PS50076; DNAJ_2; 1.
KW Chaperone; Heat shock; Repeat.
SQ SEQUENCE 446 AA; 48934 MW; 18ED5D100D5EFD3E CRC64;

Query Match 100.0%; Score 30; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYVH 5
Db 327 DTYVH 331

RESULT 31
Q8TA83 Q8TA83 PRELIMINARY; PRT; 456 AA.
AC Q8TA83;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Dnaj domain (Prokaryotic heat shock protein) protein 10, isoform
DE a.

```

Name=dnj-10; ORFNames=F22B7.5;  
 Caenorhabditis elegans.  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 Rhabditidae; Peloderinae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=Bristol N2;  
 MEDLINE=99089613; PubMed=9851916;  
 WormBase Consortium;  
 RG  
 RT "Genome sequence of the nematode *C. elegans*: a platform for  
 RT investigating biology. The *C. elegans* Sequencing Consortium. ";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Anderson K.;  
 RT "The sequence of *C. elegans* cosmid F22B7. ";  
 RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RG WormBase Consortium;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RG WormBase Consortium;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; LI2018; AAK68342.1; -.  
 DR HSSP; P08622; IXLB.  
 DR InterAct; Q8TA83; -.  
 DR WormBase; WEGene00001028; F22B7.5.  
 DR WormPep; F22B7.5a; CE24911.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0051082; F:unfolded protein binding; IEA.  
 DR GO; GO:0006457; P:protein folding; IEA.  
 DR GO; GO:0006986; P:response to unfolded protein; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR002939; DnaJ\_C.  
 DR InterPro; IPR001305; DnaJ\_CXXCXXGXG.  
 DR InterPro; IPR001623; DnaJ\_N.  
 DR InterPro; IPR008971; HSP40\_DnaJ\_pep.  
 DR InterPro; IPR003095; Hep\_DnaJ.  
 DR InterPro; IPR001878; Znf\_CCHC.  
 DR Pfam; PF00226; DnaJ; 1.  
 DR Pfam; PF01556; DnaJ\_C; 1.  
 DR Pfam; PF00684; DnaJ\_CXXCXXGXG; 1.  
 DR PRINTS; PR00625; DNAJPROTEIN.  
 DR SMART; SM00271; DnaJ; 1.  
 DR SMART; SM00343; Znf\_C2HC; 2.  
 DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
 DR PROSITE; PS00636; DnaJ\_1; 1.  
 DR PROSITE; PSS0076; DnaJ\_2; 1.  
 KW Chapterone; Heat shock; Repeat.  
 SQ SEQUENCE 456 AA; 50136 MW; 94168108B65B73B4 CRC64;

```

Qy      1 DTVH 5
      |||||
Db      327 DTVH 331

      Nature 413497-502(2002)..
      DR      EMBL; AL646080; CAD17920.1; -.
      KW      Complete proteome.
      SQ      SEQUENCE 497 AA; 55024 MW; 84625765291B492D CRC64;

Query Match      100.0%; Score 30; DB 2; Length 497;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY 1 DTYVH 5
DB 423 DTYVH 427

RESULT 34
Q935V4
ID Q935V4 PRELIMINARY; PRT; 543 AA.
AC Q935V4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pentachlorophenol-4-monoxygenase.
GN Name=pcpB;
OS Novosphingobium sp. MT1.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Novosphingobium.
OX NCBI_TaxID=145287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MT1;
RX MEDLINE=21633813; PubMed=11772624; DOI=10.1128/AEM.69.1.173-180.2002;
RA Tirola M.A., Mannisto M.K., Puhakka J.A., Kulomaa M.S.;
RT "Isolation and characterization of Novosphingobium sp. strain MT1, a
RT dominant polychlorophenol-degrading strain in a groundwater
RT bioremediation system.";
RL Appl. Environ. Microbiol. 68:173-180(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT1;
RA Tirola M.A., Wang H., Paulin L., Kulomaa M.S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ319678; CAC41015.3; -.
DR GO; GO:0015036; F:disulfide oxidoreductase activity; IEA.
DR GO; GO:0004497; F:monoxygenase activity; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR007075; Adrmdx_reductase.
DR InterPro; IPR001327; FAD pyr redox.
DR InterPro; IPR000733; Flav_monoxygenase.
DR InterPro; IPR002938; Moxv_FAD_binding.
DR InterPro; IPR000205; NAD_BS.
DR InterPro; IPR003042; Rng_mnoxygenase.
DR Pfam; PF01494; FAD binding 3; 1.
DR Pfam; PF01360; Monoxygenase; 1.
DR PRINTS; PR00419; ADXRDTASE.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00420; RNMNOXGNASE.
KW Monoxygenase.
SQ SEQUENCE 543 AA; 60341 MW; 2F9C5DF6268602D CRC64;

Query Match 100.0%; Score 30; DB 2; Length 543;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTYVH 5
DB 515 DTYVH 519

RESULT 35
Q71JB5
ID Q71JB5 PRELIMINARY; PRT; 552 AA.
AC Q71JB5
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FP13191.
GN Name=FP13191;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wan D.F., Qin W.X., Zhou X.M., Zhang P.P., Jiang H.Q., Huang Y.,
RA Zhao X.T., Gu J.R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF495722; AAQ06676.1; -.
KW Hypothetical protein.
SQ SEQUENCE 552 AA; 61681 MW; F6C8BABB46702550 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 552;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTYVH 5
DB 287 DTYVH 291

RESULT 36
Q9GV07
ID Q9GV07 PRELIMINARY; PRT; 573 AA.
AC Q9GV07
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vasa-related protein PLVAS1 (Fragment).
GN Name=Plvas1;
OS Dugesia dorotocephala.
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesidae; Dugesia.
OX NCBI_TaxID=135777;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21359115; PubMed=11466525;
RA Mochizuki K., Nishimiya-Fujisawa C., Fujisawa T.;
RT "Universal occurrence of the vasa-related genes among metazoans and
RT their germline expression in Hydra.";
RL Dev. Genes Evol. 211:299-308(2001).
CC -|- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; AB047388; BAB1313.1; -.
DR HSP; Q58083; LHV8.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase.
FT NON TER 1
SQ SEQUENCE 573 AA; 64052 MW; 6EBCE769C6C5F059 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 573;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTYVH 5
DB 452 DTYVH 456

RESULT 37
Q9Y7A2
ID Q9Y7A2 PRELIMINARY; PRT; 630 AA.
AC Q9Y7A2
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
```

01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 Mating-type protein beta 1.  
 Coprinus cinereus (Inky cap fungus).  
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 Agaricales; Pezizomycotina; Coprinopsids.  
 NCBI\_TaxID=5346;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99334617; PubMed=10406114;  
 RA Badrane H., May G.;  
 RT "The divergence-homogenization duality in the evolution of the b1  
 mating type gene of Coprinus cinereus.";  
 RL Mol. Biol. Evol. 16:975-986(1999).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99362730; PubMed=10430915; DOI=10.1073/pnas.96.16.9172;  
 RA May G., Shaw F., Badrane H., Vekemans X.;  
 RT "The signature of balancing selection: fungal mating compatibility  
 gene evolution.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9172-9177(1999).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 DR EMBL; AF126783; AAD33323.1; -.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; P:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR008422; Coprinus\_mating.  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR009057; Homeodomain like.  
 DR Pfam; PF05920; Coprinus mating; I.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; Hox; 1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 DR DNA-binding; Homeobox; Nuclear protein.  
 KW DNA-binding; Homeobox; Nuclear protein.  
 SQ SEQUENCE 630 AA; 69473 MW; A4C7AB83A7450A96 CRC64;  
 Query Match 100.0%; Score 30; DB 2; Length 630;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DTYVH 5  
 DB 51 DTYVH 55

01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 Mating-type protein beta 1.  
 Coprinus cinereus (Inky cap fungus).  
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 Agaricales; Pezizomycotina; Coprinopsids.  
 NCBI\_TaxID=5346;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99334617; PubMed=10406114;  
 RA Badrane H., May G.;  
 RT "The divergence-homogenization duality in the evolution of the b1  
 mating type gene of Coprinus cinereus.";  
 RL Mol. Biol. Evol. 16:975-986(1999).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99362730; PubMed=10430915; DOI=10.1073/pnas.96.16.9172;  
 RA May G., Shaw F., Badrane H., Vekemans X.;  
 RT "The signature of balancing selection: fungal mating compatibility  
 gene evolution.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9172-9177(1999).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 DR EMBL; AF126783; AAD33323.1; -.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; P:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR008422; Coprinus\_mating.  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR009057; Homeodomain like.  
 DR Pfam; PF05920; Coprinus mating; I.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; Hox; 1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 DR DNA-binding; Homeobox; Nuclear protein.  
 KW DNA-binding; Homeobox; Nuclear protein.  
 SQ SEQUENCE 630 AA; 69473 MW; A4C7AB83A7450A96 CRC64;  
 Query Match 100.0%; Score 30; DB 2; Length 630;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DTYVH 5  
 DB 51 DTYVH 55

01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 Mating-type protein beta 1.  
 Coprinus cinereus (Inky cap fungus).  
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 Agaricales; Pezizomycotina; Coprinopsids.  
 NCBI\_TaxID=5346;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99334617; PubMed=10406114;  
 RA Badrane H., May G.;  
 RT "The divergence-homogenization duality in the evolution of the b1  
 mating type gene of Coprinus cinereus.";  
 RL Mol. Biol. Evol. 16:975-986(1999).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99362730; PubMed=10430915; DOI=10.1073/pnas.96.16.9172;  
 RA May G., Shaw F., Badrane H., Vekemans X.;  
 RT "The signature of balancing selection: fungal mating compatibility  
 gene evolution.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9172-9177(1999).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 DR EMBL; AF126783; AAD33323.1; -.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; P:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR008422; Coprinus\_mating.  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR009057; Homeodomain like.  
 DR Pfam; PF05920; Coprinus mating; I.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; Hox; 1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 DR DNA-binding; Homeobox; Nuclear protein.  
 KW DNA-binding; Homeobox; Nuclear protein.  
 SQ SEQUENCE 630 AA; 69473 MW; A4C7AB83A7450A96 CRC64;  
 Query Match 100.0%; Score 30; DB 2; Length 630;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DTYVH 5  
 DB 51 DTYVH 55

Copper: Oxidoreductase; TPQ.  
 SQ SEQUENCE 673 AA; 75615 MW; 5F012D2A8B3DFA55 CRC64;  
 Query Match 100.0%; Score 30; DB 2; Length 673;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DTYVH 5  
 DB 352 DTYVH 356

RESULT 39  
 AMO\_PEA STANDARD; PRT; 674 AA.  
 AC Q43077;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Amine oxidase [copper-containing] precursor (EC 1.4.3.6).  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
 ON NCBI\_TaxID=3888;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Tipping A.J., McPherson M.J.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RC TISSUE=Seedling;  
 RX MEDLINE=96398683; PubMed=8805580; DOI=10.1016/S0969-2126(96)00101-3;  
 RA Kumar V., Dooley D.M., Freeman H.C., Guss J.M., Harvey I.,  
 RA McGuire M.A., Wilce M.C., Zubak V.M.;  
 RT "Crystal structure of a eukaryotic (pea seedling) copper-containing  
 amine oxidase at 2.2-A resolution.";  
 RL Structure 4:943-955(1996).  
 CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +  
 H(2)O(2).  
 CC -1- COPFACTOR: Binds 1 copper ion, 1 manganese ion and 1 topaquinone  
 per subunit.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- PTM: Topaquinone (TPQ) is generated by copper-dependent  
 autooxidation of a specific tyrosyl residue (By similarity).  
 CC -1- SIMILARITY: Belongs to the copper/topaquinone oxidase family.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; L39931; AAA62490.1; -.  
 PIR; A57327; C44239.  
 DR PDB; 1KSI; X-ray; A/B=31-672.  
 DR InterPro; IPR000269; CuNH\_oxidase.  
 DR Pfam; PF01179; Cu\_amine\_oxid; 1.  
 DR Pfam; PF02727; Cu\_amine\_oxid2; 1.  
 DR Pfam; PF02728; Cu\_amine\_oxid3; 1.  
 DR PROSITE; PS01164; COPPER AMINE OXID 1; 1.  
 DR PROSITE; PS01165; COPPER AMINE OXID 2; FALSE NEG.  
 DR 3D-structure; Copper; Glycoprotein; Manganese; Metal-binding;  
 Oxidoreductase; Signal; TPQ.  
 KW SIGNAL 1 25 Potential.  
 FT CHAIN 26 674 Amine oxidase [copper-containing].  
 FT METAL 476 476 Manganese.  
 FT METAL 477 477 Manganese (via carbonyl oxygen).  
 FT METAL 478 478 Manganese.  
 FT METAL 479 479 Manganese.  
 FT METAL 617 617 Manganese (via carbonyl oxygen).  
 FT METAL 618 618 Manganese (via carbonyl oxygen).

```

FT MOD_RES 412 2',4',5'-topaquinone.
FT METAL 467 Copper.
FT METAL 469 Copper.
FT METAL 628 Copper.
FT DISULFID 162
FT DISULFID 344
FT CARBOHYD 156
FT CARBOHYD 389
FT CARBOHYD 583
FT TURN 34
FT HELIX 40
FT TURN 56
FT TURN 58
FT STRAND 61
FT HELIX 73
FT STRAND 83
FT STRAND 91
FT TURN 99
FT STRAND 101
FT TURN 109
FT STRAND 113
FT STRAND 125
FT HELIX 130
FT TURN 137
FT HELIX 139
FT HELIX 144
FT TURN 153
FT HELIX 157
FT STRAND 160
FT STRAND 165
FT HELIX 178
FT STRAND 192
FT STRAND 196
FT TURN 200
FT STRAND 206
FT STRAND 210
FT TURN 226
FT TURN 233
FT STRAND 246
FT TURN 251
FT STRAND 257
FT TURN 260
FT STRAND 262
FT TURN 266
FT STRAND 268
FT TURN 276
FT STRAND 279
FT TURN 290
FT STRAND 294
FT TURN 316
FT STRAND 322
FT HELIX 325
FT TURN 331
FT STRAND 335
FT STRAND 340
FT TURN 341
FT STRAND 344
FT TURN 346
FT STRAND 349
FT TURN 358
FT STRAND 362
FT TURN 388
FT STRAND 392
FT TURN 410
FT STRAND 412
FT TURN 422
FT STRAND 426
FT STRAND 438
FT HELIX 446
FT STRAND 455
FT TURN 460
FT STRAND 462
FT STRAND 467
FT TURN 477
FT STRAND 480
FT STRAND 484
412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000
FT STRAND 513
FT STRAND 516
FT HELIX 517
FT STRAND 522
FT TURN 524
FT STRAND 529
FT TURN 541
FT STRAND 545
FT TURN 562
FT HELIX 565
FT TURN 570
FT HELIX 571
FT STRAND 577
FT TURN 584
FT TURN 589
FT STRAND 595
FT TURN 596
FT HELIX 603
FT HELIX 607
FT STRAND 612
FT STRAND 616
FT STRAND 618
FT HELIX 632
FT STRAND 641
FT TURN 646
FT STRAND 657
FT STRAND 666
SQ SEQUENCE 674 AA; 76358 MW; 30735390071DD18E CRC64;
Query Match 100.0%; Score 30; DB 1; Length 674;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DTVVH 5
Db 352 DTVVH 356
RESULT 40
EFGL SYNY3 STANDARD; PRT; 695 AA.
AC P28371; P74556;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Elongation factor G 1 (EF-G 1).
GN Name=fusA; Synonyms=fus; OrderedLocusNames=slr1463;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94339485; PubMed=8061323;
RA Welsch P.L., Johnson D., Zhang Y., Breitenberger C.A.;
RT "Synechocystis sp. PCC6803 fusB gene, located outside of the str
RT operon, encodes a polypeptide related to protein synthesis factor EF-
RT G.";
RL Plant Mol. Biol. 25:735-738(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugita M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -I- FUNCTION: This protein promotes the GTP-dependent translocation of
CC the nascent protein chain from the A-site to the P-site of the
CC ribosome.
```

```
CC  -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC  EF-G/EF-2 subfamily.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X65159; CAA46277.1; -.
DR  EMBL; D90916; BAA18663.1; -.
DR  PIR; S76751; S76751.
DR  HSSP; PI3551; 2EFG.
DR  HAMAP; MF_00054; -; 1.
DR  InterPro; IPR004540; EF-G.
DR  InterPro; IPR000640; EFG C.
DR  InterPro; IPR009022; EFG_III_V.
DR  InterPro; IPR005517; EFG_IV.
DR  InterPro; IPR004161; EFTU_D2.
DR  InterPro; IPR000795; ProtSyn_GTPbind.
DR  InterPro; IPR005225; Small_GTP.
DR  InterPro; IPR009000; Translat_factor.
DR  Pfam; PF00679; EFG_C; 1.
DR  Pfam; PF03764; EFG_IV; 1.
DR  Pfam; PF00009; GTP_EFTU; 1.
DR  Pfam; PF03144; GTP_EFTU_D2; 1.
DR  PRINTS; PR00315; ELONGATNPCT.
DR  TIGRFAMs; TIGR00484; EF-G; 1.
DR  TIGRFAMs; TIGR00231; small_GTP; 1.
DR  PROSITE; PS00301; EFATOR_GTP; 1.
KW  Complete proteome; Elongation factor; GTP-binding; Multigene family;
KW  Protein biosynthesis.
FT  NP_BIND 15 22 GTP (By similarity).
FT  NP_BIND 79 83 GTP (By similarity).
FT  NP_BIND 133 136 GTP (By similarity).
FT  CONFLICT 42 42 H -> R (in Ref. 1).
SQ  SEQUENCE 695 AA; 76750 MW; 8727830BF9073223 CRC64;

Query Match          100.0%; Score 30; DB 1; Length 695;
Best Local Similarity 100.0%; Pred.No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTYYH 5
Db 490 DTYYH 494
```

Search completed: June 3, 2005, 16:01:29  
Job time : 45.4242 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 15:37:35 ; Search time 156.606 Seconds  
(without alignments)  
41.984 Million cell updates/sec

Title: US-10-089-452-22

Perfect score: 92

Sequence: 1 KIDPANGTKYDPIFQA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_l6Dec04.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	17	4 AAB86082	Aab86082 H. pylori
2	92	100.0	17	4 AAB86050	Aab86050 H. pylori
3	92	100.0	120	4 AAB86107	Aab86107 H. pylori
4	92	100.0	120	4 AAB86069	Aab86069 H. pylori
5	73	79.3	17	5 ABB75622	Abb75622 Anti-RANK
6	73	79.3	17	8 ADS94368	Ads94368 Antibody
7	73	79.3	93	7 ABU62850	Abu62850 Mouse ant
8	73	79.3	113	2 AAR60527	Aar60527 Heavy cha
9	73	79.3	119	6 ABP97650	Abp97650 Amino aci
10	73	79.3	120	2 AAR79889	Aar79889 Anti-EGFR
11	73	79.3	125	8 ADR40435	Adr40435 Mouse 2c
12	73	79.3	125	8 ADR40353	Adr40353 Mouse 2c
13	73	79.3	133	5 ABB75617	Abb75617 Anti-RANK
14	73	79.3	136	2 AAR41679	Aar41679 V heavy c
15	73	79.3	141	8 ADS94356	Ads94356 Antibody
16	73	79.3	437	2 AAW37738	Aaw37738 Nucleotid
17	72	78.3	110	7 ABU62845	Abu62845 Mouse ant
18	71	77.2	121	8 ADR40415	Adr40415 Reshaped
19	71	77.2	121	8 ADR40333	Adr40333 Reshaped
20	71	77.2	123	2 AAR81330	Aar81330 Mouse ant
21	71	77.2	123	2 AAR81323	Aar81323 Humanized
22	71	77.2	123	2 AAW22413	Aaw22413 Humanized
23	71	77.2	123	8 ADR40413	Adr40413 Mouse 21.
24	71	77.2	123	8 ADR40424	Adr40424 Mouse 21.
25	71	77.2	123	8 ADR40426	Adr40426 Reshaped

26	71	77.2	123	8	ADR40331	Adr40331 Mouse 21.
27	71	77.2	123	8	ADR40342	Adr40342 Mouse 21.
28	71	77.2	123	8	ADR40344	Adr40344 Reshaped
29	71	77.2	140	2	AAR81327	Aar81327 Mouse VLA
30	71	77.2	140	2	AAW22410	Aaw22410 Alpha-4 i
31	71	77.2	142	2	AAR81333	Aar81333 Human VLA
32	71	77.2	142	2	AAW22428	Aaw22428 Humanised
33	71	77.2	142	8	ADR40432	Adr40432 Reshaped
34	71	77.2	142	8	ADR40350	Adr40350 Reshaped
35	71	77.2	144	8	ADR40419	Adr40419 Mouse 21.
36	71	77.2	144	8	ADR40337	Adr40337 Mouse 21.
37	69	75.0	17	7	ABR61875	AbR61875 Mouse Mab
38	69	75.0	17	8	ADQ31264	Adq31264 CDR2 pept
39	69	75.0	117	7	ABR61872	AbR61872 Mouse Mab
40	69	75.0	117	8	ADQ31281	Adq31281 Humanised
41	69	75.0	117	8	ADQ31261	Adq31261 Murine 11
42	69	75.0	117	8	ADQ31282	Adq31282 Humanised
43	69	75.0	206	6	ABP96767	Abp96767 TSH recep
44	69	75.0	206	6	ABP96768	Abp96768 TSH recep
45	69	75.0	254	2	AAy30117	Aay30117 Murine an

#### ALIGNMENTS

##### RESULT 1

AAB86082  
ID AAB86082 standard; peptide; 17 AA.

XX AAB86082;

DT 17-JUN-2001 (first entry)

DE H. pylori catalase derived antibody HP25/6m/1BS heavy chain CDR2.

KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;  
KW acid-resistant microorganism; complementarity determining region; CDR;  
KW feces; heavy chain; light chain.

XX Unidentified.

OS WO200127612-A2.

PN 19-APR-2001.

PD 12-OCT-2000; 2000WO-EP010057.

XX 12-OCT-1999; 99EP-00120351.

PR 16-MAR-2000; 2000EP-00105592.

PR 31-MAR-2000; 2000EP-00107028.

PR 10-MAY-2000; 2000EP-00110110.

XX (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

PI Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;

XX WPI; 2001-282086/29.

XX N-ESDB; AAF88109.

DR Detecting infections by acid-resistant microorganisms, particularly for  
diagnosing Helicobacter pylori, comprises immunochromatographic detection  
of antigen in feces.

XX Claim 26; Page 25; 90pp; German.

CC This invention describes a novel method for detecting infection by an  
acid-resistant microorganism (A), in a mammal, using  
immunochromatography. The method is used to diagnose infection by an acid  
-resistant microorganism (A), in a mammal, such as Helicobacter,  
Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,  
inexpensive and non-invasive, and may indicate the stage of infection. A  
test strip used in the method may include a filter to eliminate particles

CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence represents a  
 CC complementarity determining region (CDR) from an antibody raised against  
 CC the H. pylori catalase or beta-urease antigen which is used to illustrate  
 CC the method of the invention

XX SQ Sequence 17 AA;

Query Match 100.0%; Score 92; DB 4; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17  
 |||||  
 Db 1 KIDPANGTKYDPIFOA 17

# RESULT 2

AAB86050  
 ID AAB86050 standard; peptide; 17 AA.

XX AC AAB86050;

XX DT 17-JUL-2001 (first entry)

XX DE H. pylori catalase derived antibody HP25/6m/1B5 heavy chain CDR2.

XX KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
 KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
 KW antibacterial; complementarity determining region.

XX OS Unidentified.

XX PN WO200127613-A2.

XX PD 19-APR-2001.

XX PF 12-OCT-2000; 2000WO-EP010058.

XX PR 12-OCT-1999; 99EP-00120351.

XX PR 16-MAR-2000; 2000EP-00105592.

XX PR 31-MAR-2000; 2000EP-00107028.

XX PR 10-MAY-2000; 2000EP-00110110.

XX PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX PI Reiter C, Cullmann G, Heppner P, Ringels A, Mueller H, Haindl E;  
 XX WPI; 2001-282087/29.

XX DR N-PSDB; AAF88052.

XX PT Detecting infections by acid-resistant microorganisms, particularly for  
 XX diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
 XX sample.

XX PS Claim 19; Page 16; 89pp; German.

XX CC This invention describes a novel method for detecting, in a mammal,  
 CC infection by an acid-resistant microorganism (A) which comprises reacting  
 CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
 CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
 CC complex is formed with Ag, and the formation of a complex detected. R are  
 CC specific for an Ag which, after passage through the intestines, at least  
 CC in some mammals, retains a native (or corresponding) structure against  
 CC which the mammal produces antibodies (when immunized or infected with  
 CC (A), or its extracts, lysates or derived proteins (or fragments) or  
 CC synthetic peptides). The products of the invention have antibacterial  
 CC activity. The method is used to diagnose infection by Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the  
 CC progress of treatment. Receptors, particularly antibodies, directed

CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an  
 CC infection. This sequence represents a complementarity determining region  
 CC (CDR) from an antibody generated against a Helicobacter pylori antigen  
 CC (catalase or beta-urease) which is used to illustrate the method of the  
 CC invention

XX SQ Sequence 17 AA;

Query Match 100.0%; Score 92; DB 4; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17  
 |||||  
 Db 1 KIDPANGTKYDPIFOA 17

# RESULT 3

AAB86107

ID AAB86107 standard; protein; 120 AA.

XX AC AAB86107;

XX DT 17-JUL-2001 (first entry)

XX DE H. pylori catalase antibody HP25/2m/1B5 V region heavy chain.

XX KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;  
 KW acid-resistant microorganism; complementarity determining region; CDR;  
 KW feces; heavy chain; light chain.

XX OS Unidentified.

XX PN WO200127612-A2.

XX PD 19-APR-2001.

XX PF 12-OCT-2000; 2000WO-EP010057.

XX PR 12-OCT-1999; 99EP-00120351.

XX PR 16-MAR-2000; 2000EP-00105592.

XX PR 31-MAR-2000; 2000EP-00107028.

XX PR 10-MAY-2000; 2000EP-00110110.

XX PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX PI Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;  
 XX WPI; 2001-282086/29.

XX DR N-PSDB; AAF88153.

XX PT Detecting infections by acid-resistant microorganisms, particularly for  
 XX diagnosing Helicobacter pylori, comprises immunochromatographic detection  
 XX of antigen in feces.

XX PS Claim 31; Fig 3; 90pp; German.

XX CC This invention describes a novel method for detecting infection by an  
 CC acid-resistant microorganism (A), in a mammal, using  
 CC immunochromatography. The method is used to diagnose infection by an acid  
 CC -resistant microorganism (A), in a mammal, such as Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,  
 CC inexpensive and non-invasive, and may indicate the stage of infection. A  
 CC test strip used in the method may include a filter to eliminate particles  
 CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence represents a

CC Helicobacter pylori catalase' derived antibody V-region heavy chain  
 CC fragment used to illustrate the method of the invention  
 XX  
 SQ Sequence 120 AA;

Query Match 100.0%; Score 92; DB 4; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-07;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQA 17  
 |||||  
 Db 50 KIDPANGTKYDPIFQA 66

RESULT 4  
 AAB86069  
 ID AAB86069 standard; protein; 120 AA.  
 XX AC AAB86069;  
 XX  
 DT 17-JUL-2001 (first entry)  
 XX  
 DE H. pylori catalase antibody HP25/6m/1B5 V-region heavy chain.  
 XX  
 KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
 KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
 KW antibacterial; complementarity determining region; V-region.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200127613-A2.  
 XX  
 PD 19-APR-2001.  
 XX  
 PF 12-OCT-2000; 2000WO-EP010058.  
 XX  
 PR 12-OCT-1999; 99EP-00120351.  
 PR 16-MAR-2000; 2000EP-00105592.  
 PR 31-MAR-2000; 2000EP-00107028.  
 PR 10-MAY-2000; 2000EP-00110110.  
 XX  
 PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
 XX  
 PI Reiter C, Cullmann G, Heppner P, Ringeis A, Mueller H, Haindl E;  
 XX  
 DR WPI; 2001-282087/29.  
 DR N-PSDB; AAF88096.  
 XX  
 PT Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
 PT sample.  
 XX  
 PS Claim 27; Fig 3; 89pp; German.  
 XX

This invention describes a novel method for detecting, in a mammal,  
 CC infection by an acid-resistant microorganism (A) which comprises reacting  
 CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
 CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
 CC complex is formed with Ag, and the formation of a complex detected. R are  
 CC specific for an Ag which, after passage through the intestines, at least  
 CC in some mammals, retains a native (or corresponding) structure against  
 CC which the mammal produces antibodies (when immunized or infected with  
 CC (A), or its extracts, lysates or derived proteins (or fragments) or  
 CC synthetic peptides). The products of the invention have antibacterial  
 CC activity. The method is used to diagnose infection by Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the  
 CC progress of treatment. Receptors, particularly antibodies, directed  
 CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an

CC infection. This sequence represents a Helicobacter pylori anti-catalase  
 CC derived antibody HP25/6m/1B5 V-region heavy chain fragment which is  
 CC described in the method of the invention  
 XX  
 SQ Sequence 120 AA;

Query Match 100.0%; Score 92; DB 4; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-07;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQA 17  
 |||||  
 Db 50 KIDPANGTKYDPIFQA 66

RESULT 5  
 ABB75622  
 ID ABB75622 standard; peptide; 17 AA.  
 XX AC ABB75622;  
 XX  
 DT 10-JUN-2002 (first entry)  
 XX  
 DE Anti-RANK ligand mAb 19H22 heavy chain variable region CDR2.  
 XX  
 KW RANK ligand; RANK-L; monoclonal antibody; mAb; 19H22; antibody;  
 KW rheumatoid arthritis; osteoporosis; bone cancer; metastasis; osteolysis;  
 KW osteoarthritis; psoriasis; diabetes; multiple sclerosis;  
 KW inflammatory bowel disease; cytostatic; antiarthritic; antiinflammatory;  
 KW antidiabetic; osteopathic; antirheumatic; immunosuppressive;  
 KW antipsoriatic; neuroprotective; heavy chain; HL; mouse.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200215846-A2.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 21-AUG-2001; 2001WO-US026161.  
 PR 21-AUG-2000; 2000US-0226524P.  
 PR 07-SEP-2000; 2000US-0230639P.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX  
 PI Sweet RW, Tornetta MA, Truneh A, Wattam TA;  
 XX  
 DR WPI; 2002-242022/29.  
 DR N-PSDB; ABL75622.  
 XX  
 PT A mouse monoclonal antibody 19H22 for treating or preventing osteopenic  
 PT disease including rheumatoid arthritis and osteoporosis.  
 XX  
 PS Claim 5; Page 43; 45pp; English.  
 XX

The present sequence is that of complementarity determining region 2  
 CC (CDR2) of the heavy chain variable region of claimed anti-human RANK  
 CC ligand (RANK-L) murine monoclonal antibody (mAb) 19H22 (isotype  
 CC IgG2b/kappa). The mAb is produced by hybridoma cells obtained by  
 CC immunising CB6 f1 mice with human RANK-L, and fusing spleen cells with  
 CC X63-Ag8.653 myeloma cells. 19H22 mAb is specific for human RANK-L, having  
 CC a binding affinity characterised by a dissociation constant of about 10  
 CC key regulator of the immune system, bone development and homeostasis. The  
 CC invention provides altered, chimeric and humanised 19H22 antibodies,  
 CC including Fab or F(ab')2 fragments, antibodies comprising the light and  
 CC heavy chain variable regions of 19H22, and antibodies comprising the CDRs  
 CC of 19H22. These are used in methods for the diagnosis, treatment and  
 CC prevention of osteopenic diseases, including rheumatoid arthritis,  
 CC osteoporosis, metastatic and primary bone cancer, wear debris induced  
 CC osteolysis or osteoarthritis, and immune diseases including psoriasis,  
 CC insulin dependent diabetes, inflammatory bowel disease or multiple

CC sclerosis. Methods are also provided for the recombinant production of  
CC the antibodies, using isolated nucleic acids and transfected host  
CC (preferably mammalian) host cells, and a method of diagnosing conditions  
CC associated with Th1 T-cell activity or osteoclast development and  
CC activation, in particular those listed above  
XX  
XX Sequence 17 AA;

Query Match 79.3%; Score 73; DB 5; Length 17;  
Best Local Similarity 81.2%; Pred. No. 0.00011;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQFQ 16  
:||||| ||||| ||  
Db 1 RIDPANGNTKYDPKRFQ 16

## RESULT 6

ADS94368  
ID ADS94368 standard; peptide; 17 AA.

XX  
AC ADS94368;

XX  
DT 02-DEC-2004 (first entry)

XX  
DE Antibody 16H9 VH2 CDR peptide SEQ ID NO:67.

XX antibody; antigen-binding antibody fragment;  
KW cell-associated CA 125/O772P; monoclonal antibody; cytostatic;  
KW immunostimulant; mediator of lysis; tumour; cell proliferative disorder;  
KW cancer; cervical cancer; uterine cancer; breast cancer; lung cancer;  
KW ovarian cancer.  
XX

OS Synthetic.

XX  
PN WO2004035537-A2.

XX  
PD 29-APR-2004.

XX  
PF 15-OCT-2003; 2003WO-US032945.

XX  
PR 16-OCT-2002; 2002US-0418828P.

PR 10-JUL-2003; 2003US-0485986P.

XX  
PA (EURO-) EUROCELTIQUE SA.

XX  
PI Albione EF, Soltis DA;

XX  
DR WPI; 2004-357171/33.

XX  
PT Novel isolated antibody, or antigen-binding antibody fragment binding  
PT with cell-associated CA 125/O772P polypeptide relative to shed CA  
PT 125/O772 polypeptide, useful for ameliorating cervical or ovarian cancer.

XX  
PS Disclosure; SEQ ID NO 67; 153pp; English.

XX  
CC The present invention describes an isolated antibody, or an antigen-  
CC binding antibody fragment (I), that preferentially binds cell-associated  
CC CA 125/O772P polypeptide relative to shed CA 125/O772 polypeptide. Also  
CC described: (1) a monoclonal antibody (II) produced by hybridoma 4E7; (2)  
CC a monoclonal antibody that competes with binding of (II); (3) a hybridoma  
CC as deposited in (II); (4) an isolated nucleic acid molecule (III)  
CC comprising a nucleotide sequence that encodes a variable chain region of  
CC (I); (5) a pharmaceutical composition comprising an antibody or an  
CC antigen-binding antibody fragment that preferentially binds cell-  
CC associated CA 125/O772P polypeptide relative to shed CA 125/O772P  
CC polypeptide, and a carrier; (6) a pharmaceutical composition comprising a  
CC monoclonal antibody or an antigen-binding monoclonal antibody fragment  
CC that preferentially binds cell-associated CA 125/O772P polypeptide  
CC relative to shed CA 125/O772P polypeptide, and a carrier; (7) an article  
CC of manufacture (IV) comprising packaging material and a composition  
CC comprising an antibody, or an antigen-binding antibody fragment that  
CC preferentially binds cell-associated CA 125/O772P relative to shed CA

CC 125/O772P, and a carrier contained within the packaging material, and  
CC composition in a form suitable for administration to a subject; (8) a  
CC fusion polypeptide (V) comprising an antibody, or an antigen-binding  
CC antibody fragment, which preferentially binds cell-associated CA  
CC 125/O772P relative to shed CA 125/O772P operably linked to a heterologous  
CC agent; (9) ameliorating (M1) a symptom of a CA 125/O772P-related disorder  
CC; (10) a monoclonal antibody (VI) chosen from 325.1, 621.1, 633.1, 654.1,  
CC 725.1, 8G9, 7F10, 8A1, 8AC3, 15C9, 8E3, 8B5, 7G10, 16C7, 7C6, 7H1, 16H9,  
CC 7A11, 4E7, 117.1, 368.1, 446.1, 501.1, and 776.1 or its antigen-binding  
CC antibody fragment; (11) an antibody or antigen binding antibody fragment  
CC that competes with (VI); and (12) a pharmaceutical composition comprising  
CC (VI) and a carrier. (I) has cytostatic activity, and can be used as an  
CC immunostimulant and a mediator of lysis of positive tumour cell. (I) is  
CC useful for ameliorating a symptom of a CA 125/O772P-related disorder which  
CC is a cell proliferative disorder such as cancer, cervical or uterine  
CC cancer, breast or lung cancer or ovarian cancer. (V) is useful  
CC diagnostically for monitoring the development or progression of cancer or  
CC tumour as part of clinical testing procedure. The present sequence  
CC represents an antibody complementarity determining region (CDR) amino  
CC acid sequence, which is used in the exemplification of the present  
CC invention.  
XX  
XX Sequence 17 AA;

Query Match 79.3%; Score 73; DB 8; Length 17;  
Best Local Similarity 81.2%; Pred. No. 0.00011;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQFQ 16  
:||||| ||||| ||  
Db 1 RIDPANGNTKYDPKRFQ 16

## RESULT 7

ABU62850

ID ABU62850 standard; protein; 93 AA.

XX  
AC ABU62850;

XX  
DT 17-SEP-2003 (first entry)

XX  
DE Mouse anti-LPS monoclonal antibody VH H10.

XX  
KW WaaP; tyrosine kinase; enzyme; lipopolysaccharide; LPS; HepI; mouse;  
KW antibody; heavy chain variable region; VH H10; monoclonal antibody;  
KW inner core oligosaccharide; phosphorylation; bacterial infection; VH 7-4.

XX  
OS Mus sp.

XX  
PN US2003017518-A1.

XX  
PD 23-JAN-2003.

XX  
PF 26-JUN-2002; 2002US-00179851.

XX  
PR 26-JUN-2001; 2001US-0300420P.

XX  
PA (LAMJ/) LAM J.  
PA (ZHAO/) ZHAO X.

XX  
PI Lam J, Zhao X;

XX  
DR WPI; 2003-521592/49.  
DR N-PSDB; ACD27478.

XX  
PT Assay for modulators of enzymes involved in phosphorylation of the inner  
PT core oligosaccharide of a lipopolysaccharide, comprises detecting  
PT phosphorylated lipopolysaccharide in a test sample after adding a  
PT specific antibody.

XX  
PS Disclosure; Fig 14C; 33pp; English.

XX  
CC The invention relates to assaying for modulators of an enzyme involved in

the phosphorylation of the inner core oligosaccharide (HepI) of lipopolysaccharides (LPS), comprises: (a) incubating a test sample comprising the enzyme (e.g. the tyrosine kinase Waap), a candidate substance and substrates comprising dephosphorylated LPS and a source of phosphate; (b) adding at least one antibody that binds to phosphorylated LPS while not binding to dephosphorylated LPS; and (c) detecting phosphorylated LPS in the test sample by measuring the binding of the antibody to phosphorylated LPS (by an ELISA assay, enzyme linked immunosorbent assay). An increase or decrease in the amount of phosphorylated LPS in the test sample in the presence of the candidate substance indicates that the candidate substance is a modulator. Also included is a kit for performing the method. The method is for assaying for modulators, preferably inhibitors, of an enzyme involved in the phosphorylation of the inner core oligosaccharide of LPS, wherein a decrease in the amount of phosphorylated LPS in the test sample in the presence of the candidate substance indicates that the candidate substance is an inhibitor. It is useful in conducting a target discovery business by providing assay systems using the method of the invention for identifying agents by their ability to modulate and enzyme involved in the phosphorylation of the inner core oligosaccharide of LPS, conducting therapeutic profiling of identified agents for efficacy and toxicity in animals, and licensing the rights to a third party for further drug development and/or sales or the identified agents or their analogues. The method does not require the use of radio-labeled substrates. It is amenable to automation. LPS a major virulence factor for infectious bacteria such as *P. aeruginosa*. The present sequence represents the heavy chain variable region from mouse monoclonal antibody H10, the parental *aeruginosa* LPS

Query Match 79.3%; Score 73; DB 7; Length 93;  
Best Local Similarity 81.2%; Pred. No. 0.00068;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16  
:|||||:|||||:  
Db 45 RIDPANGTKYDPKRFQ 60

Sequence 93 AA;

RESULT 8  
AAR60527  
ID AAR60527 standard; protein; 113 AA.  
AC AAR60527;  
XX  
XX  
DT 25-MAR-2003 (revised)  
DT 07-APR-1995 (first entry)  
XX  
XX  
DE Heavy chain variable region of monoclonal antibody 5E8.  
XX  
XX Secretion sequence; vector; antibody; single chain antibody; promoter;  
XX Bacillus subtilis.  
XX  
XX Synthetic.  
XX  
XX EP614982-A1.  
XX  
XX 14-SEP-1994.  
XX  
XX 05-MAR-1994; 94EP-00200564.  
XX  
XX 10-MAR-1993; 93IT-MI000456.  
XX  
XX (ENIE ) ENIRICERCH SPA.  
XX  
XX Grandi G, De Ferra F, Tosi C, Tortora O, Cuzzoni A;  
XX WPI; 1994-281209/35.  
XX DR N-PSDB; AAQ71285.  
XX  
XX Vector for secretion of single chain antibody from *Bacillus subtilis* -

contg. neutral protease promoter and specific leader sequence, providing high yield of soluble antibody for diagnostic or therapeutic use.

Example 3; Fig 1; 27pp; English.

The sequence encoding the variable region of the heavy chain of the monoclonal antibody 5E8 (specific for the alpha subunit of human gonadotropin) was amplified for its use in a recombinant vector which also comprised (1) the promoter of the gene for the neutral protease of *Bacillus subtilis* BSC 1A341, (2) the secretion sequence described in AAQ71278. The amplified fragment was used to produce a DNA sequence encoding an antibody in single molecule form with the sequence VH/VK-L-VK/VH-(TAG)n, where VH and VK are the variable regions of the heavy and light chains of the antibody and L is the linker of sequence Val-Ser-(Gly(4)-Ser)3. TAG is a peptide recognised by polyclonal antibodies directed towards the same peptide, n is 1 or 0. The vector is used for the production of single chain antibodies which have improved pharmacokinetic properties and can be produced more economically than monoclonal antibodies. The vector allows production of antibodies in completely soluble form with secretion in high yield. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to correct PR field.)

Sequence 113 AA;

Query Match 79.3%; Score 73; DB 2; Length 113;  
Best Local Similarity 81.2%; Pred. No. 0.00084;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16  
:|||||:|||||:  
Db 50 RIDPANGTKYDPKRFQ 65

RESULT 9  
ABP97650  
ID ABP97650 standard; protein; 119 AA.  
XX  
XX AC ABP97650;  
XX  
XX DT 16-MAY-2003 (first entry)  
XX  
XX DE Amino acid sequence of K121 antibody heavy chain variable region.  
XX  
XX KW K121 antibody; K121-like antibody; kappa-type myeloma cell;  
XX kappa-type multiple myeloma; haematopoietic cell transplantation;  
XX apoptosis; kappa myeloma antigen.  
XX  
XX OS Mus musculus.  
XX  
XX FH Key Location/Qualifiers  
XX Region 26..35  
XX FT /note= "complementarity determining region"  
XX Region 50..66  
XX FT /note= "complementarity determining region"  
XX Region 99..108  
XX FT /note= "complementarity determining region"  
XX  
XX WO2003004056-A1.  
XX  
XX PN  
XX PD 16-JAN-2003.  
XX  
XX PF 05-JUL-2002; 2002WO-AU000896.  
XX  
XX PR 06-JUL-2001; 2001AU-00006179.  
XX  
XX PA (PACM-) PACMAB PTY LTD.  
XX  
XX FI Raison RL, Dunn RD, Choo BHA;  
XX WPI; 2003-210317/20.  
XX DR N-PSDB; ABZ68619.  
XX  
XX PT Treating kappa-type multiple myeloma in a subject by administering a K121

PT -like antibody not conjugated to a toxin or a cytolytic agent.

XX Example 8; Fig 9a; 65pp; English.

CC The present sequence is a murine K121 antibody heavy chain variable region. This antibody is used as the base for designing a K121-like antibody, that competes with K121 for binding to kappa-type myeloma cells. The K121-like antibody is used in the method of the invention. The specification describes a method for treating kappa-type multiple myeloma in a subject, comprising administering a K121-like antibody which is not conjugated to a toxin or a cytolytic agent. The method is useful for treating kappa-type multiple myeloma, autologous haematopoietic cell transplantation, killing kappa-type myeloma cells in a mixed population of cells and inducing apoptosis in kappa myeloma antigen (KMA) bearing cells

XX Sequence 119 AA;

SQ Query Match 79.3%; Score 73; DB 6; Length 119;  
Best Local Similarity 81.2%; Pred. No. 0.00089;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:||||| ||||| ||  
Db 50 RIDPANGTKYDPKFQ 65

RESULT 10

AAR79889  
ID AAR79889 standard; protein; 120 AA.

XX AC AAR79889;

XX DT 02-JUL-1996 (first entry)

XX DE Anti-EGFR antibody heavy chain variable region (Clone S3 12D).

XX Single chain antibody; antibody; epidermal growth factor receptor; EGFR;  
KW tumour; cancer; glioma; melanoma; carcinoma; diagnosis; assessment;  
KW phage antibody library.

XX OS Mus musculus.

XX FH Key Location/Qualifiers  
FT Region 1..30  
FT /label= Framework region 1.  
FT Region 31..35  
FT /label= CDR1.  
FT Region 36..49  
FT /label= Framework region 2.  
FT Region 50..66  
FT /label= CDR2.  
FT Region 67..98  
FT /label= Framework region 3.  
FT Region 99..109  
FT /label= CDR3.  
FT Region 110..120  
FT /label= Framework region 4.

XX WO9525167-A1.

XX PD 21-SEP-1995.

XX PF 16-MAR-1995; 95WO-EP000978.

XX PR 17-MAR-1994; 94EP-00104160.

XX PR 02-DEC-1994; 94EP-00118970.

XX PA (MERE ) MERCK PATENT GMBH.

XX Kettleborough AC, Bendig MM, Ansell KH, Guessow D, Adan J;

XX Mitjans F, Rosell E, Blasco F, Piulats J;

XX

DR WPI; 1995-336972/43.

XX Anti-EGFR antibodies and single chain Fv antibody fragments - obtained from phage-antibody libraries, useful for diagnosis and therapy of tumours.

XX Disclosure; Fig 1B; 93pp; English.

XX Anti-epidermal growth factor receptor (EGFR) single chain antibodies and antibodies constructed from anti-EGFR antibody fragments can be used for diagnosis of tumours and assessment of tumour growth in vitro and in vivo. They may also be used in a pharmaceutical composition for the therapy of e.g. melanomas, gliomas and carcinomas. The antibodies and fragments are derived from mice but are humanised so as to cause minimum reaction against them. They are produced using the phage antibody library. (See AAT04011-T04026 and AAR79858-R79873)

XX Sequence 120 AA;

SQ Query Match 79.3%; Score 73; DB 2; Length 120;  
Best Local Similarity 81.2%; Pred. No. 0.00089;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:||||| ||||| ||  
Db 50 RIDPANGTKYDPKFQ 65

RESULT 11

ADR40435  
ID ADR40435 standard; protein; 125 AA.

XX AC ADR40435;

XX DT 04-NOV-2004 (first entry)

XX DE Mouse 2c antibody Vh region amino acid fragment.

XX Sulphonylamine; remyelination; demyelinating disease;  
KW lymphocyte infiltration; paralysis; neuroprotective; muscular;  
KW cytostatic; antiinflammatory; vulneryary;  
KW selective adhesion molecule inhibitor; SAMI; antibody; 21.6; mouse.

XX OS Mus sp.

XX FH Key Location/Qualifiers  
FT Misc-difference 106  
FT /note= "residue Xaa is unknown"  
FT Misc-difference 120  
FT /note= "residue Xaa is unknown"

XX WO2004066931-A2.

XX PD 12-AUG-2004.

XX PF 26-JAN-2004; 2004WO-US002028.

XX PR 24-JAN-2003; 2003US-0442171P.

XX PR 05-SEP-2003; 2003US-0500316P.

XX PA (ELAN-) ELAN PHARM INC.

XX Karlik SJ, pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;  
PI Messersmith E, Freedman S, Yednock T;

XX WPI; 2004-593990/57.

XX Use of sulfonylamine compounds as selective adhesion molecule inhibitors to treat demyelinating disease e.g. multiple sclerosis, by inhibiting lymphocyte infiltration and promoting remyelination to treat paralysis.

XX Example 3; Fig 18A; 573pp; English.

XX

CC The invention relates to sulphonylamine compounds (A) of specified  
 CC formulae. The compounds are used to prepare a medicament to promote  
 CC remyelination of nerve cells in a mammal for treatment of a demyelinating  
 CC disease and to inhibit lymphocyte infiltration of immune cells in the  
 CC spinal cord to promote remyelination of nerve cells in the spinal cord,  
 CC thus treating paralysis. The method of inhibiting lymphocyte infiltration  
 CC of immune cells in the spinal cord to promote remyelination of nerve  
 CC cells in the spinal cord and thus treating paralysis further comprises co  
 CC -administering an immunosuppressant such as adrenocorticotrophic hormone,  
 CC cortisol, cortisone, fludrocortisone, methylprednisolone, dexamethasone  
 CC methylprednisolone, triamcinolone or betamethasone), alpha-  
 CC (e.g. interferon beta-1b or interferon beta-1a). (A) is used to prepare a  
 CC medicament for treating paralysis, a congenital metabolic disorder, a  
 CC neuropathy with abnormal myelination, drug induced demyelination,  
 CC radiation induced demyelination, hereditary demyelinating condition,  
 CC prior induced demyelinating condition, encephalitis induced demyelination  
 CC or spinal cord injury. The present sequence represents a mouse 2c  
 CC antibody Vh region fragment leading to the design of reshaped human 21.6  
 CC antibody heavy chain variable regions.

XX Sequence 125 AA;

Query Match 79.3%; Score 73; DB 8; Length 125;  
 Best Local Similarity 81.2%; Pred. No. 0.00093;  
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
 :||||| ||||| ||  
 Db 50 RIDPANGTKYDPKFK 65

RESULT 12

ADRA40353  
 ID ADRA40353 standard; protein; 125 AA.

XX AC ADRA40353;

DT 04-NOV-2004 (first entry)

XX Mouse 2c antibody Vh region amino acid fragment.

XX Remyelination; demyelinating disease; lymphocyte infiltration; paralysis;  
 KW neuroprotective; muscular; cytosolic; antiinflammatory; vulnary;  
 KW selective adhesion molecule inhibitor; SAM1; antibody; 21.6; mouse.

XX Mus sp.

PH Key Location/Qualifiers

FT Misc-difference 106 /note= "residue Xaa is unknown"

FT Misc-difference 120 /note= "residue Xaa is unknown"

XX WO2004066932-A2.

XX 12-AUG-2004.

XX 26-JAN-2004; 2004WO-US002039.

XX 24-JAN-2003; 2003US-044217P.

PR 05-SEP-2003; 2003US-0500316P.

XX (ELAN-) ELAN PHARM INC.

XX Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;

PI Messersmith E, Freedman S, Yednock T;

XX WPI; 2004-593991/57.

XX Use of remyelinating agent such as natalizumab for preparation of  
 PT medicament for treating demyelinating diseases and paralysis by promoting  
 PT remyelination of nerve cells in mammal.

XX Example 3; Fig 18A; 495pp; English.

XX The invention relates to remyelinating agents (A) that are used to  
 CC prepare a medicament to promote remyelination of nerve cells in a mammal  
 CC for treatment of a demyelinating disease and to inhibit lymphocyte  
 CC infiltration of immune cells in the spinal cord to promote remyelination  
 CC of nerve cells in the spinal cord, thus treating paralysis. The method of  
 CC inhibiting lymphocyte infiltration of immune cells in the spinal cord to  
 CC promote remyelination of nerve cells in the spinal cord and thus treating  
 CC paralysis further comprises co-administering an immunosuppressant such as  
 CC adrenocorticotrophic hormone, a corticosteroid (e.g. prednisone,  
 CC methylprednisolone, dexamethasone, cortisone, fludrocortisone,  
 CC prednisolone, alpha-methylprednisolone, triamcinolone or betamethasone),  
 CC or an interferon (e.g. interferon beta-1b or interferon beta-1a). (A) is  
 CC used to prepare a medicament for treating paralysis, a congenital  
 CC metabolic disorder, a neuropathy with abnormal myelination, drug induced  
 CC demyelination, radiation induced demyelination, hereditary demyelinating  
 CC condition, prior induced demyelinating condition, encephalitis induced  
 CC demyelination or spinal cord injury. The present sequence represents a  
 CC mouse 2c antibody Vh region fragment leading to the design of reshaped  
 CC human 21.6 antibody heavy chain variable regions.

XX Sequence 125 AA;

Query Match 79.3%; Score 73; DB 8; Length 125;  
 Best Local Similarity 81.2%; Pred. No. 0.00093;  
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
 :||||| ||||| ||  
 Db 50 RIDPANGTKYDPKFK 65

RESULT 13

ABB75617  
 ID ABB75617 standard; protein; 133 AA.

XX AC ABB75617;

XX 10-JUN-2002 (first entry)

XX Anti-RANK ligand mAb 19H22 heavy chain variable region.

XX RANK ligand; RANK-L; monoclonal antibody; mAb; 19H22; antibody;  
 KW complementarity determining region; CDR; rheumatoid arthritis;  
 KW osteoporosis; bone cancer; metastasis; osteolysis; osteoarthritis;  
 KW psoriasis; inflammatory bowel disease; multiple sclerosis; diabetes;  
 KW cytosolic; antiarthritic; antiinflammatory; osteopathic; antidiabetic;  
 KW antirheumatic; immunosuppressive; antipsoriatic; neuroprotective;  
 KW heavy chain; HL; mouse.

XX Mus musculus.

PH Key Location/Qualifiers

FT Region 31..35

FT /label= CDR1

FT /note= "complementarity determining region 1"

FT Region 50..66

FT /label= CDR2

FT /note= "complementarity determining region 2"

FT Region 99..109

FT /label= CDR3

FT /note= "complementarity determining region 3"

XX WO200215846-A2.

XX 28-FEB-2002.

XX 21-AUG-2001; 2001WO-US026161.

XX 21-AUG-2000; 2000US-0226524P.

PR 07-SEP-2000; 2000US-0230639P.

```

XX (SMIK ) SMITHKLINE BEECHAM CORP.
PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX
XX Sweet RW, Tornetta MA, Truneh A, Wattam TA;
XX
XX WPI; 2002-242022/29.
DR N-PSDB; ABL53531.
XX
XX A mouse monoclonal antibody 19H22 for treating or preventing osteopenic
PT disease including rheumatoid arthritis and osteoporosis.
PT
XX Claim 6; Page 8; 45pp; English.
XX
XX The present sequence is that of the heavy chain variable region of
CC claimed anti-human RANK ligand (RANK-L) murine monoclonal antibody (mAb)
CC 19H22 (isotype IGG2b/kappa). The mAb is produced by hybridoma cells
CC obtained by immunising CB6 f1 mice with human RANK-L protein, and fusing
CC spleen cells with X63-Ag8.653 myeloma cells. 19H22 mAb is specific for
CC human RANK-L, having a binding affinity characterised by a dissociation
CC constant of about 10 power -10 M. RANK-L is a member of the tumour
CC necrosis factor family of proteins and a key regulator of the immune
CC system, bone development and homeostasis. The invention provides altered,
CC chimeric and humanised 19H22 antibodies, including Fab or F(ab')2
CC fragments, antibodies comprising the light and heavy chain variable
CC regions of 19H22, and antibodies comprising the complementarity
CC determining regions of 19H22. These are used in methods for the
CC diagnosis, treatment and prevention of osteopenic diseases, including
CC rheumatoid arthritis, osteoporosis, metastatic and primary bone cancer,
CC wear debris induced osteolysis or osteoarthritis, and immune diseases
CC including psoriasis, insulin dependent diabetes, inflammatory bowel
CC disease or multiple sclerosis. Methods are also provided for the
CC recombinant production of the antibodies, using isolated nucleic acids
CC and transfected host (preferably mammalian) host cells, and a method of
CC diagnosing conditions associated with Th1 T-cell activity or osteoclast
CC development and activation, in particular those listed above
XX
XX Sequence 133 AA;
SQ
Query Match 79.3%; Score 73; DB 5; Length 133;
Best Local Similarity 81.2%; Pred. No. 0.001;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db :||||| ||||| ||
50 RIDPANGNTKYDPRFQ 65

RESULT 14
AAR41679
XX AAR41679 standard; protein; 136 AA.
XX
XX AAR41679;
XX
XX 25-MAR-2003 (revised)
DT 24-MAR-1994 (first entry)
XX
XX V heavy chain of recombinant anti-feline calcivirus antibody.
DE
XX Antibody; feline calcivirus; FCV; neutralise; heavy chain; virus;
KW infection; diagnosis; treatment; prophylaxis.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH 47..54
FT Domain /label= Complementary Determining Region 1.
FT 69..86
FT Domain /label= Complementary Determining Region 2.
FT 118..125
FT Domain /label= Complementary Determining Region 3.
XX
XX EP561194-A2.
PN

XX 22-SEP-1993.
PD
XX 26-FEB-1993; 93EP-00103066.
PF
XX 28-FEB-1992; 92JP-00079189.
PR
XX (KAGA ) CHEMA SERO THERAPEUTIC RES INST.
PA
XX Kimachi K, Maeda H, Nishiyama K, Tokiyoshi S, Tohya Y, Mikami T;
PI WPI; 1993-296521/38.
XX DR N-PSDB; AAQ48001.
DR
XX Anti-feline calcivirus recombinant antibody - used to treat, prevent and
PT diagnosis infection and is safe but effective in cats.
PT
XX Claim 1-2; Fig 3; 36pp; English.
XX
XX The gene fragment encoding the V heavy chain region is used to produce a
CC genetically engineered antibody capable of reacting specifically with
CC feline calcivirus (FCV). The antibody can be used to treat, prevent and
CC diagnose FCV infection. It has a broad virus- neutralising spectrum and
CC can be used to neutralise viruses which have acquired immunity to other
CC neutralising antibodies. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 136 AA;
SQ
Query Match 79.3%; Score 73; DB 2; Length 136;
Best Local Similarity 81.2%; Pred. No. 0.001;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db :||||| ||||| ||
69 RIDPANGNTKYDPRFQ 84

RESULT 15
ADS94356
ID ADS94356 standard; protein; 141 AA.
XX
XX ADS94356;
AC
XX 02-DEC-2004 (first entry)
DT
XX Antibody 16H9 heavy chain variable region (16H9H) SEQ ID NO:55.
DE
XX antibody; antigen-binding antibody fragment;
KW cell-associated CA 125/O772p; monoclonal antibody; cytostatic;
KW immunostimulant; mediator of lysis; tumour; cell proliferative disorder;
KW cancer; cervical cancer; uterine cancer; breast cancer; lung cancer;
KW ovarian cancer.
XX
XX Synthetic.
OS
XX WO2004035537-A2.
XX PN
XX 29-APR-2004.
XX PD
XX 15-OCT-2003; 2003WO-US032945.
PF
XX 16-OCT-2002; 2002US-0418828P.
PR
XX 10-JUL-2003; 2003US-0485986P.
PR
XX (EURO-) EUROCELTIQUE SA.
XX PA
XX Albone EF, Soltis DA;
XX PI WPI; 2004-357171/33.
XX DR N-PSDB; ADS94359.
DR
XX Novel isolated antibody, or antigen-binding antibody fragment binding
PT with cell-associated CA 125/O772p polypeptide relative to shed CA
PN

```



125/O772 polypeptide, useful for ameliorating cervical or ovarian cancer.  
 Claim 40; SEQ ID NO 55; 153pp; English.

The present invention describes an isolated antibody, or an antigen-binding antibody fragment (I), that preferentially binds cell-associated CA 125/O772P polypeptide relative to shed CA 125/O772 polypeptide. Also described: (1) a monoclonal antibody (II) produced by hybridoma 4E7; (2) a monoclonal antibody that competes with binding of (II); (3) a hybridoma as deposited in (II); (4) an isolated nucleic acid molecule (III) comprising a nucleotide sequence that encodes a variable chain region of (I); (5) a pharmaceutical composition comprising an antibody or an antigen-binding antibody fragment that preferentially binds cell-associated CA 125/O772P polypeptide relative to shed CA 125/O772P polypeptide, and a carrier; (6) a pharmaceutical composition comprising a monoclonal antibody or an antigen-binding monoclonal antibody fragment that preferentially binds cell-associated CA 125/O772P polypeptide relative to shed CA 125/O772P polypeptide, and a carrier; (7) an article of manufacture (IV) comprising packaging material and a composition comprising an antibody, or an antigen-binding antibody fragment that preferentially binds cell-associated CA 125/O772P relative to shed CA 125/O772P, and a carrier contained within the packaging material, and a composition in a form suitable for administration to a subject; (8) a fusion polypeptide (V) comprising an antibody, or an antigen-binding antibody fragment, which preferentially binds cell-associated CA 125/O772P relative to shed CA 125/O772P operably linked to a heterologous agent; (9) ameliorating (MI) a symptom of a CA 125/O772P-related disorder; (10) a monoclonal antibody (VI) chosen from 325.1, 621.1, 633.1, 654.1, 725.1, 8G9, 7F10, 8A1, 8AC3, 15C9, 8E3, 8B5, 7G10, 16C7, 7C6, 7H1, 16H9, 7A11, 4E7, 117.1, 368.1, 446.1, 501.1, and 776.1 or its antigen-binding antibody fragment; (11) an antibody or antigen binding antibody fragment that competes with (VI); and (12) a pharmaceutical composition comprising (VI) and a carrier. (I) has cytostatic activity, and can be used as an immunostimulant and a mediator of lysis of positive tumour cell. (I) is useful for ameliorating a symptom of a CA 125/O772P-related disorder which is a cell proliferative disorder such as cancer, cervical or uterine cancer, breast or lung cancer or ovarian cancer. (V) is useful diagnostically for monitoring the development or progression of cancer or tumour as part of clinical testing procedure. The present sequence represents an antibody heavy chain variable region amino acid sequence, which is used in the exemplification of the present invention.

Sequence 141 AA;

Query Match 79.3%; Score 73; DB 8; Length 141;  
 Best Local Similarity 81.2%; Pred. No. 0.0011;  
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
 :||||| ||||| ||  
 Db 69 RIDPANGNTKYDPKFKQ 84

RESULT 16

AAW37738

ID AAW37738 standard; protein; 437 AA.

XX

AC AAW37738;

XX

DT 17-OCT-2003 (revised)

DT 07-JUL-1998 (first entry)

XX Nucleotide sequence encoding the Mgr6-clavin immunotoxin.

DE Recombinant ribosomal inhibitor protein; RIP; clavin; inhibition;

XX protein synthesis; conjugate; Mgr6-clavin; anticancer; antiviral agent.

XX Aspergillus clavatus.

OS Mus musculus.

OS Chimeric.

XX Key

XX Location/Qualifiers

Peptide 96..286

FT Peptide /note= "Mgr6 from M. musculus"  
 FT 287..437  
 FT /note= "Clavin from A. clavatus"

XX WO9749726-A1.

PN 31-DEC-1997.

XX 26-JUN-1997; 97WO-EP003359.

XX 27-JUN-1996; 96IT-PI000155.

XX (ITU-) ITAL MIN UNIV RICERCA SCI & TECNOLOGICA.

XX Mele A, De Santis R, Parente D, Colnaghi M;

XX WPI; 1998-077109/07.

XX N-PSDB; AAV09256.

XX DNA encoding Aspergillus clavatus-derived ribosomal inhibitor protein,

XX clavin - useful as an immuno-conjugate and for treatment of cancer.

XX Claim 4; Page 14-15; 26pp; English.

XX This is the amino acid sequence of a fusion protein comprising of the Mgr6

XX single chain antibody (with the heavy and light chain joined by a

XX linker), fused to the clavin protein. Clavin is an inhibitor protein, and

XX functions by inactivating the ribosomes. Clavin or its conjugates are

XX useful as anticancer and/or antiviral agents. The recombinant ribosomal

XX inhibitor protein (RIP), Mgr6-clavin (a conjugate of clavin) is used as

XX an immunoconjugate, and the complex and clavin alone are capable of

XX inhibiting ribosomal activity. (Updated on 17-OCT-2003 to standardise OS

XX field)

XX Sequence 437 AA;

Query Match 79.3%; Score 73; DB 2; Length 437;

Best Local Similarity 81.2%; Pred. No. 0.0035;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16

:||||| ||||| ||

Db 86 RIDPANGNTKYDPKFKQ 101

RESULT 17

ABU62845

ID ABU62845 standard; protein; 110 AA.

XX

AC ABU62845;

XX

DT 17-SEP-2003 (first entry)

XX Mouse anti-LPS monoclonal antibody VH 7-4.

XX WaaP; tyrosine kinase; enzyme; lipopolysaccharide; LPS; HepI; mouse;

XX antibody; heavy chain variable region; VH 7-4; monoclonal antibody;

XX inner core oligosaccharide; phosphorylation; bacterial infection.

XX Mus sp.

XX US2003017518-A1.

XX 23-JAN-2003.

XX 26-JUN-2002; 2002US-00179851.

XX 26-JUN-2001; 2001US-0300420P.

XX (LAMJ/) LAM J.

XX (ZHAO/) ZHAO X.

XX Lam J, Zhao X;



XX PA (ELAN-) ELAN PHARM INC.

XX PI Karlik SJ, Pleiss WA, Konradi AW, Grant FS, Semko CM, Dresseen D;

XX PI Messeremith E, Freedman S, Yednock T;

XX DR WPI; 2004-593991/57.

XX PI Use of remyelinating agent such as natalizumab for preparation of

PT medicament for treating demyelinating diseases and paralysis by promoting

PT remyelination of nerve cells in mammal.

XX Example 3; Page 449-452; 495pp; English.

XX PS The invention relates to remyelinating agents (A) that are used to

CC prepare a medicament to promote remyelination of nerve cells in a mammal

CC for treatment of a demyelinating disease and to inhibit lymphocyte

CC infiltration of immune cells in the spinal cord to promote remyelination

CC of nerve cells in the spinal cord, thus treating paralysis. The method of

CC inhibiting lymphocyte infiltration of immune cells in the spinal cord to

CC promote remyelination of nerve cells in the spinal cord and thus treating

CC paralysis further comprises co-administering an immunosuppressant such as

CC adrenocorticotrophic hormone, a corticosteroid (e.g. prednisone,

CC prednisolone, alpha-methylprednisolone, triamcinolone or betamethasone)

CC or an interferon (e.g. interferon beta-1b or interferon beta-1a). (A) is

CC used to prepare a medicament for treating paralysis, a congenital

CC metabolic disorder, a neuropathy with abnormal myelination, drug induced

CC demyelination, radiation induced demyelination, hereditary demyelinating

CC condition, prion induced demyelinating condition, encephalitis induced

CC demyelination or spinal cord injury. The present sequence represents a

CC reshaped human 21.6 antibody heavy chain variable region fragment.

XX Sequence 121 AA;

Query Match 77.2%; Score 71; DB 8; Length 121;

Best Local Similarity 81.2%; Pred. No. 0.0019;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16

:||||| ||||| ||

Db 50 RIDPANGTKYDKPKFQ 65

RESULT 20

AAR81330

ID AAR81330 standard; protein; 123 AA.

XX AC AAR81330;

XX DT 02-APR-1996 (first entry)

XX DE Mouse anti-VLA-4 antibody 21.6 heavy chain variable region.

XX KW Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;

XX KW antibody engineering.

XX OS Mus musculus.

XX FH Key Location/Qualifiers

XX FT 1..30

FT /label= FR1

FT /note= "mouse heavy chain variable framework region 1"

FT 31..35

FT /label= CDR1

FT /note= "mouse heavy chain variable complementarity

FT determining region 1"

FT 36..49

FT /label= FR2

FT /note= "mouse heavy chain variable framework region 2"

FT 50..66

FT /label= CDR2

FT /note= "mouse heavy chain variable complementarity

determining region 2"

67..98

/label= FR3

/note= "mouse heavy chain variable framework region 3"

99..112

/label= CDR3

/note= "mouse heavy chain variable complementarity

determining region 3"

113..123

/label= FR4

/note= "mouse heavy light chain variable framework region

4"

W09519790-A1.

XX 27-JUL-1995.

XX 25-JAN-1995; 95WO-US001219.

XX 25-JAN-1994; 94US-00186269.

XX (ATHE-) ATHENA NEUROSCIENCES INC.

XX Bendig MM, Leger OJ, Saldanha J, Jones TS;

XX WPI; 1995-269276/35.

XX New humanised antibodies against VLA-4 - used for inhibiting leukocyte

adhesion to endothelial cells, partic. for treating inflammatory disease.

XX Disclosure; Page 68; 105pp; English.

XX The sequence represents the mouse anti-VLA-4 antibody 21.6 heavy chain

variable region (without signal sequence). Cloned cDNA CDR sequences of

mouse 21.6 variable light and variable heavy regions are linked to human

constant framework regions of the REI antibody for the light chain and

the 2\*CL antibody for the heavy chain in the construction of a humanized

antibody against VLA-4. The 5' and 3' ends of the mouse CDNA are

modified using PCR primers (See AAQ99895-98) and then subcloned into

mammalian cell expression vectors containing human kappa or gamma-1

constant regions. In the humanized light chain, amino acids L45, L49, L58

and L69 in the human kappa LCVR framework are replaced by the amino acid

present in the equivalent position of the mouse 21.6 Ig light chain.

Plasmids encoding the chimeric antibodies are transfected into COS cells.

The humanized antibodies can be used to inhibit adhesion of a leukocyte

to an endothelial cell and to treat inflammatory diseases such as

multiple sclerosis. They can also be used in the treatment of stroke,

cerebral traumas, meningitis or encephalitis. The antibodies can also be

used for detecting VLA-4, for affinity purification or for generating

anti-idiotypic antibodies

XX Sequence 123 AA;

Query Match 77.2%; Score 71; DB 2; Length 123;

Best Local Similarity 81.2%; Pred. No. 0.0019;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16

:||||| ||||| ||

Db 50 RIDPANGTKYDKPKFQ 65

RESULT 21

AAR81323

ID AAR81323 standard; protein; 123 AA.

XX AC AAR81323;

XX DT 02-APR-1996 (first entry)

XX DE Humanized VLA-4 antibody 21.6 heavy chain variable region, Ha.

XX KW Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;

```

KW antibody engineering.
XX Mus musculus.
XX WO9519790-A1.
XX 27-JUL-1995.
XX
XX 25-JAN-1995; 95WO-US001219.
XX
XX 25-JAN-1994; 94US-00186269.
XX (ATHE-) ATHENA NEUROSCIENCES INC.
XX
XX Bendig MM, Leger OJ, Saldanha J, Jones TS;
XX WPI; 1995-269276/35.
XX
XX New humanised antibodies against VLA-4 - used for inhibiting leukocyte
XX adhesion to endothelial cells, partic. for treating inflammatory disease.
XX
XX Claim 11; Page 69; 105pp; English.
XX
XX The sequence encodes the mouse antibody 21.6 heavy chain variable region,
XX Ha, directed against leukocyte adhesion molecule VLA-4. Cloned cDNA
XX sequences of mouse 21.6 VL and VH (AAQ99889 and AAQ99892) regions are
XX linked to human constant regions in the construction of a humanized
XX antibody against VLA-4. The 5' and 3' ends of the mouse cDNAs are
XX modified using PCR primers (See AAQ99895-98) and then subcloned into
XX mammalian cell expression vectors containing human kappa or gamma-1
XX constant regions. In the humanized light chain, amino acids L45, L49, L58
XX and L69 in the human kappa LC VR framework are replaced by the amino acid
XX present in the equivalent position of the mouse 21.6 Ig L chain. Plasmids
XX encoding the chimeric antibodies are transfected into COS cells. The
XX humanized antibodies can be used for inhibiting adhesion of a leukocyte
XX to an endothelial cell and for treating inflammatory diseases such as
XX multiple sclerosis. They can also be used in the treatment of stroke,
XX cerebral traumas, meningitis or encephalitis. The antibodies can also be
XX used for detecting VLA-4, for affinity purification or for generating
XX anti-idiotypic antibodies
XX
XX Sequence 123 AA;

Query Match 77.2%; Score 71; DB 2; Length 123;
Best Local Similarity 81.2%; Pred. No. 0.0019;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFQ 16
Db 50 RIDPANGTKYDKPKFQ 65

RESULT 22
AAW22413
ID AAW22413 standard; protein; 123 AA.
XX
XX AAW22413;
XX
XX 08-DEC-1997 (first entry)
XX
XX Humanised alpha-4 integrin antibody 21.6 VH Ha.
XX
XX Alpha-4 integrin; humanised antibody; monoclonal antibody 21.6; asthma;
XX atherosclerosis; AIDS; dementia; diabetes; tumour; metastasis;
XX inflammatory bowel disease; rheumatoid arthritis; transplant rejection;
XX graft versus host disease; nephritis; atopic dermatitis; psoriasis;
XX myocardial ischaemia; acute leukocyte mediated lung injury; therapy.
XX
XX Mus; musculus.
XX Homo; sapiens.
XX Synthetic.
XX Chimeric.
XX

```

---

```

FH Key Location/Qualifiers
FT 1. .30
FT /label= FR1
FT /note= "21/28'CL framework region 1"
FT Misc-difference 27. .30
FT /note= "21/28'CL residues 27-30 are replaced by those of
FT MAB 21.6, involved in antigen binding"
FT 31. .35
FT /label= CDR1
FT /note= "21.6 complementarity determining region 1"
FT 36. .49
FT /label= FR2
FT /note= "21/28'CL framework region 2"
FT 50. .66
FT /label= CDR2
FT /note= "21.6 complementarity determining region 2"
FT 67. .98
FT /label= FR3
FT /note= "21/28'CL framework region 3"
FT Misc-difference 72
FT /note= "21/28'CL Arg-72 is substd. by Ala of mouse 21.6
FT VL, important in supporting the CDR2 loop"
FT 99. .112
FT /label= CDR3
FT /note= "21.6 complementarity determining region 3"
FT 113. .123
FT /label= FR4
FT /note= "21/28'CL framework region 4"
XX
XX WO9718838-A1.
XX
XX 29-MAY-1997.
XX
XX 21-NOV-1996; 96WO-US018807.
XX
XX 21-NOV-1995; 95US-00561521.
XX (ATHE-) ATHENA NEUROSCIENCES INC.
XX
XX Bendig MM, Leger OJ, Saldanha J, Jones ST, Yednock TA;
XX WPI; 1997-297879/27.
XX
XX Uses of humanised alpha-4 integrin antibody - for treatment of asthma,
XX atherosclerosis, AIDS, dementia, etc.
XX
XX Claim 25; Fig 7; 107pp; English.
XX
XX This polypeptide, designated Ha, comprises the heavy chain variable
XX region (VH) of a humanised alpha-4 integrin antibody 21.6. It is composed
XX of complementarity determining regions (CDRs) from the VH region (see
XX AAW22410) of mouse alpha-4 integrin monoclonal antibody 21.6 and a
XX modified human 21/28'CL framework. It can be expressed in mammalian host
XX cells following PCR amplification and mutagenesis of appropriate
XX fragments of mouse and human DNA sequences. The humanised 21.6 VH and a
XX humanised 21.6 VL (see AAW22412) can be used to produce a claimed
XX humanised 21.6 antibody that is useful in the manufacture of a medicament
XX for treating asthma, atherosclerosis, AIDS, dementia, diabetes,
XX inflammatory bowel disease, rheumatoid arthritis, transplant rejection,
XX graft versus host disease, tumour metastasis, nephritis, atopic
XX dermatitis, psoriasis, myocardial ischaemia, and acute leukocyte mediated
XX lung injury. The antibody may also be used in the affinity purification
XX of alpha-4 integrin for use as a vaccine or an immunogen. It is also
XX useful for generating idiotypic antibodies. The humanised antibody has a
XX half-life in the human circulation essentially equivalent to that of
XX naturally occurring human antibodies
XX
XX Sequence 123 AA;

Query Match 77.2%; Score 71; DB 2; Length 123;
Best Local Similarity 81.2%; Pred. No. 0.0019;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

Oy 1 KIDPANGTKYDPIFQ 16  
:|||||

Db 50 RIDPANGTKYDPKFQ 65

RESULT 23  
ADR40413  
ID ADR40413 standard; protein; 123 AA.  
AC ADR40413;  
XX 04-NOV-2004 (first entry)  
XX Mouse 21.6 antibody Vh region amino acid fragment.  
XX Sulphonylamine; remyelination; demyelinating disease;  
KW lymphocyte infiltration; paralysis; neuroprotective; muscular;  
KW cyostatic; antiinflammatory; vulnerary;  
KW selective adhesion molecule inhibitor; SAMI; antibody; 21.6; mouse.  
XX Mus sp.  
XX WO2004066931-A2.  
XX 12-AUG-2004.  
XX 26-JAN-2004; 2004WO-US002028.  
XX 24-JAN-2003; 2003US-0442171P.  
XX 05-SEP-2003; 2003US-0500316P.  
XX (ELAN-) ELAN PHARM INC.  
XX Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;  
XX Messersmith E, Freedman S, Yednock T;  
XX WPI; 2004-593990/57.  
XX Use of sulfonamide compounds as selective adhesion molecule inhibitors  
XX to treat demyelinating disease e.g. multiple sclerosis, by inhibiting  
XX lymphocyte infiltration and promoting remyelination to treat paralysis.  
XX Example 3; Page 449-452; 573pp; English.  
XX The invention relates to sulphonylamine compounds (A) of specified  
XX formulae. The compounds are used to prepare a medicament to promote  
XX remyelination of nerve cells in a mammal for treatment of a demyelinating  
XX disease and to inhibit lymphocyte infiltration of immune cells in the  
XX spinal cord to promote remyelination of nerve cells in the spinal cord,  
XX thus treating paralysis. The method of inhibiting lymphocyte infiltration  
XX of immune cells in the spinal cord to promote remyelination of nerve  
XX cells in the spinal cord and thus treating paralysis further comprises co  
XX -administering an immunosuppressant such as adrenocorticotrophic hormone,  
XX a corticosteroid (e.g. prednisone, methylprednisolone, dexamethasone,  
XX cortisol, cortisone, fludrocortisone, prednisolone, alpha-  
XX methylprednisolone, triamcinolone or betamethasone) or an interferon  
XX (e.g. interferon beta-1b or interferon beta-1a). (A) is used to prepare a  
XX medicament for treating paralysis, a congenital metabolic disorder, a  
XX neuropathy with abnormal myelination, drug induced demyelination,  
XX radiation induced demyelination, hereditary demyelinating condition,  
XX prior induced demyelinating condition, encephalitis induced demyelination  
XX or spinal cord injury. The present sequence represents a mouse 21.6  
XX antibody Vh region fragment leading to the design of reshaped human 21.6  
XX antibody heavy chain variable regions.

SQ Sequence 123 AA;

Query Match 77.2%; Score 71; DB 8; Length 123;  
Best Local Similarity 81.2%; Pred. No. 0.0019;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KIDPANGTKYDPIFQ 16  
:|||||

Db 50 RIDPANGTKYDPKFQ 65

RESULT 24  
ADR40424  
ID ADR40424 standard; protein; 123 AA.  
AC ADR40424;  
XX 04-NOV-2004 (first entry)  
XX Mouse 21.6 antibody Vh region amino acid fragment.  
XX Sulphonylamine; remyelination; demyelinating disease;  
KW lymphocyte infiltration; paralysis; neuroprotective; muscular;  
KW cyostatic; antiinflammatory; vulnerary;  
KW selective adhesion molecule inhibitor; SAMI; antibody; 21.6; mouse.  
XX Mus sp.  
XX WO2004066931-A2.  
XX 12-AUG-2004.  
XX 26-JAN-2004; 2004WO-US002028.  
XX 24-JAN-2003; 2003US-0442171P.  
XX 05-SEP-2003; 2003US-0500316P.  
XX (ELAN-) ELAN PHARM INC.  
XX Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;  
XX Messersmith E, Freedman S, Yednock T;  
XX WPI; 2004-593990/57.  
XX Use of sulfonamide compounds as selective adhesion molecule inhibitors  
XX to treat demyelinating disease e.g. multiple sclerosis, by inhibiting  
XX lymphocyte infiltration and promoting remyelination to treat paralysis.  
XX Example 3; Fig 14; 573pp; English.  
XX The invention relates to sulphonylamine compounds (A) of specified  
XX formulae. The compounds are used to prepare a medicament to promote  
XX remyelination of nerve cells in a mammal for treatment of a demyelinating  
XX disease and to inhibit lymphocyte infiltration of immune cells in the  
XX spinal cord to promote remyelination of nerve cells in the spinal cord,  
XX thus treating paralysis. The method of inhibiting lymphocyte infiltration  
XX of immune cells in the spinal cord to promote remyelination of nerve  
XX cells in the spinal cord and thus treating paralysis further comprises co  
XX -administering an immunosuppressant such as adrenocorticotrophic hormone,  
XX a corticosteroid (e.g. prednisone, methylprednisolone, dexamethasone,  
XX cortisol, cortisone, fludrocortisone, prednisolone, alpha-  
XX methylprednisolone, triamcinolone or betamethasone) or an interferon  
XX (e.g. interferon beta-1b or interferon beta-1a). (A) is used to prepare a  
XX medicament for treating paralysis, a congenital metabolic disorder, a  
XX neuropathy with abnormal myelination, drug induced demyelination,  
XX radiation induced demyelination, hereditary demyelinating condition,  
XX prior induced demyelinating condition, encephalitis induced demyelination  
XX or spinal cord injury. The present sequence represents a mouse 21.6  
XX antibody Vh region fragment.

SQ Sequence 123 AA;

Query Match 77.2%; Score 71; DB 8; Length 123;  
Best Local Similarity 81.2%; Pred. No. 0.0019;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KIDPANGTKYDPIFQ 16  
:|||||

Db 50 RIDPANGTKYDPKFQ 65

XX	RESULT 25
AC	ADR40426
ID	ADR40426 standard; protein; 123 AA.
XX	
AC	ADR40426;
XX	
DT	04-NOV-2004 (first entry)
XX	
DE	Reshaped human 21.6 antibody Vh region fragment Ha.
XX	
KW	Sulphonylamine; remyelination; demyelinating disease;
KW	lymphocyte infiltration; paralysis; neuroprotective;
KW	cytostatic; antiinflammatory; muscular;
XX	selective adhesion molecule inhibitor; SAMI; antibody; 21.6; human.
XX	
OS	Homo sapiens.
XX	
PN	WO2004066931-A2.
XX	
PD	12-AUG-2004.
XX	
PY	26-JAN-2004; 2004WO-US002028.
XX	
PR	24-JAN-2003; 2003US-0442171P.
XX	
PR	05-SEP-2003; 2003US-0500316P.
XX	(ELAN-) ELAN PHARM INC.
PA	
PI	Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;
PI	Messersmith E, Freedman S, Yednock T;
XX	
DR	WPI; 2004-593990/57.
XX	
PT	Use of sulfonylamine compounds as selective adhesion molecule inhibitors
PT	to treat demyelinating disease e.g. multiple sclerosis, by inhibiting
PT	lymphocyte infiltration and promoting remyelination to treat paralysis.
XX	
PS	Example 3; Fig 14; 573pp; English.
XX	
CC	The invention relates to sulphonylamine compounds (A) of specified
CC	formulae. The compounds are used to prepare a medicament to promote
CC	remyelination of nerve cells in a mammal for treatment of a demyelinating
CC	disease and to inhibit lymphocyte infiltration of immune cells in the
CC	spinal cord to promote remyelination of nerve cells in the spinal cord,
CC	thus treating paralysis. The method of inhibiting lymphocyte infiltration
CC	of immune cells in the spinal cord to promote remyelination of nerve
CC	cells in the spinal cord and thus treating paralysis further comprises co
CC	-administering an immunosuppressant such as adrenocorticotrophic hormone,
CC	a corticosteroid (e.g. prednisone, methylprednisolone, dexamethasone
CC	cortisol, cortisone, fludrocortisone, prednisolone, alpha-
CC	methylprednisolone, triamcinolone or betamethasone) or an interferon
CC	(e.g. interferon beta-lb or interferon beta-la). (A) is used to prepare a
CC	medicament for treating paralysis, a congenital metabolic disorder, a
CC	neuropathy with abnormal myelination, drug induced demyelination,
CC	radiation induced demyelination, hereditary demyelinating condition,
CC	prior induced demyelinating condition, encephalitis induced demyelination
CC	or spinal cord injury. The present sequence represents a version of the
CC	reshaped human 21.6 antibody heavy chain variable region fragment.
XX	
SQ	Sequence 123 AA;
	Query Match 77.2%; Score 71; DB 8; Length 123;
	Best Local Similarity 81.2%; Pred. No. 0.0019;
	Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY	1 KIDPANGKTKYDPFIQ 16
	:
Dd	50 RIDPANGYTKYDKPKFQ 65
	RESULT 26
AD	ADR40331
DD	ADR40331 standard; protein; 123 AA.

```

XX DE Mouse 21.6 antibody Vh region amino acid fragment.
XX DE
XX DE Remyelination; demyelinating disease; lymphocyte infiltration; paralysis;
XX DE neuroprotective; muscular; cytosolic; antiinflammatory; vulnery;
XX DE selective adhesion molecule inhibitor; SAMI; antibody; 21.6; mouse.
XX OS
XX OS Mus sp.
XX PN WO2004066932-A2.
XX PD 12-AUG-2004.
XX PF 26-JAN-2004; 2004WO-US002039.
XX PR 24-JAN-2003; 2003US-0442171P.
XX PR 05-SEP-2003; 2003US-0500316P.
XX PA (ELAN-) ELAN PHARM INC.
XX PI Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;
XX PI Messersmith E, Freedman S, Yednock T;
XX DR WPI; 2004-593991/57.
XX PS
XX PS Use of remyelinating agent such as natalizumab for preparation of
XX PT medicament for treating demyelinating diseases and paralysis by promoting
XX PT remyelination of nerve cells in mammal.
XX PS
XX PS Example 3; Fig 14; 495pp; English.
XX CC The invention relates to remyelinating agents (A) that are used to
XX CC prepare a medicament to promote remyelination of nerve cells in a mammal
XX CC for treatment of a demyelinating disease and to inhibit lymphocyte
XX CC infiltration of immune cells in the spinal cord to promote remyelination
XX CC of nerve cells in the spinal cord, thus treating paralysis. The method of
XX CC inhibiting lymphocyte infiltration of immune cells in the spinal cord to
XX CC promote remyelination of nerve cells in the spinal cord and thus treating
XX CC paralysis further comprises co-administering an immunosuppressant such as
XX CC adrenocorticotrophic hormone, a corticosteroid (e.g. prednisone,
XX CC methylprednisolone, dexamethasone cortisone, triamcinolone or betamethasone)
XX CC or an interferon (e.g. interferon beta-1b or interferon beta-1a). (A) is
XX CC used to prepare a medicament for treating paralysis, a congenital
XX CC metabolic disorder, a neuropathy with abnormal myelination, drug induced
XX CC demyelination, radiation induced demyelination, hereditary demyelinating
XX CC condition, prion induced demyelinating condition, encephalitis induced
XX CC demyelination or spinal cord injury. The present sequence represents a
XX CC mouse 21.6 antibody Vh region fragment.
XX SQ Sequence 123 AA;

Query Match 77.2%; Score 71; DB 8; Length 123;
Best Local Similarity 81.2%; Pred. No. 0.0019;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16
Db :||||| ||||| ||
50 RIDPANGTKYDPKPFQ 65

RESULT 28
ADR40344
ID ADR40344 standard; protein; 123 AA.
XX AC
XX AC ADR40344;
XX DT
XX DT 04-NOV-2004 (first entry)
XX DE Reshaped human 21.6 antibody Vh region fragment Ha.
XX DE Remyelination; demyelinating disease; lymphocyte infiltration; paralysis;
XX DE neuroprotective; muscular; cytosolic; antiinflammatory; vulnery;
XX KW

XX KW Mouse 21.6 antibody Vh region amino acid fragment.
XX OS
XX OS Homo sapiens.
XX PN WO2004066932-A2.
XX PD 12-AUG-2004.
XX PF 26-JAN-2004; 2004WO-US002039.
XX PR 24-JAN-2003; 2003US-0442171P.
XX PR 05-SEP-2003; 2003US-0500316P.
XX PA (ELAN-) ELAN PHARM INC.
XX PI Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;
XX PI Messersmith E, Freedman S, Yednock T;
XX DR WPI; 2004-593991/57.
XX PS
XX PS Use of remyelinating agent such as natalizumab for preparation of
XX PT medicament for treating demyelinating diseases and paralysis by promoting
XX PT remyelination of nerve cells in mammal.
XX PS
XX PS Example 3; Fig 14; 495pp; English.
XX CC The invention relates to remyelinating agents (A) that are used to
XX CC prepare a medicament to promote remyelination of nerve cells in a mammal
XX CC for treatment of a demyelinating disease and to inhibit lymphocyte
XX CC infiltration of immune cells in the spinal cord to promote remyelination
XX CC of nerve cells in the spinal cord, thus treating paralysis. The method of
XX CC inhibiting lymphocyte infiltration of immune cells in the spinal cord to
XX CC promote remyelination of nerve cells in the spinal cord and thus treating
XX CC paralysis further comprises co-administering an immunosuppressant such as
XX CC adrenocorticotrophic hormone, a corticosteroid (e.g. prednisone,
XX CC methylprednisolone, dexamethasone cortisone, triamcinolone or betamethasone)
XX CC or an interferon (e.g. interferon beta-1b or interferon beta-1a). (A) is
XX CC used to prepare a medicament for treating paralysis, a congenital
XX CC metabolic disorder, a neuropathy with abnormal myelination, drug induced
XX CC demyelination, radiation induced demyelination, hereditary demyelinating
XX CC condition, prion induced demyelinating condition, encephalitis induced
XX CC demyelination or spinal cord injury. The present sequence represents a
XX CC mouse 21.6 antibody Vh region fragment.
XX SQ Sequence 123 AA;

Query Match 77.2%; Score 71; DB 8; Length 123;
Best Local Similarity 81.2%; Pred. No. 0.0019;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16
Db :||||| ||||| ||
50 RIDPANGTKYDPKPFQ 65

RESULT 29
AAR81327
ID AAR81327 standard; protein; 140 AA.
XX AC
XX AC AAR81327;
XX DT
XX DT 23-MAR-1996 (first entry)
XX DE Mouse VLA-4 antibody 21.6 light heavy variable region.
XX DE Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;
XX DE antibody engineering.
XX OS Mus musculus.
XX PH Key Location/Qualifiers

```

```

FT Peptide      1. .19
FT /note= "signal peptide"
FT Region      20. .49
FT /note= "framework region 1"
FT Region      50. .54
FT /note= "complementarity determining region 1"
FT Region      55. .68
FT /note= "framework region 2"
FT Region      69. .85
FT /note= "complementarity determining region 2"
FT Region      86. .117
FT /note= "framework region 3"
FT Region      118. .131
FT /note= "complementarity determining region 3"
FT Region      132. .140
FT /note= "framework region 4"
FT XX
XX WO9519790-A1.
XX
XX 27-JUL-1995.
XX
XX 25-JAN-1995; 95WO-US001219.
XX
XX 25-JAN-1994; 94US-00186269.
XX
XX (ATHE-) ATHENA NEUROSCIENCES INC.
XX
XX Bendig MM, Leger OJ, Saldanha J, Jones TS;
XX
XX WPI; 1995-269276/35.
XX N-PSDB; AAQ99892.
XX
XX New humanised antibodies against VLA-4 - used for inhibiting leukocyte
XX adhesion to endothelial cells, partic. for treating inflammatory disease.
XX
XX Disclosure; Fig 2; 105pp; English.
XX
XX The sequence represents the mouse antibody 21.6 heavy chain variable
XX region directed against leukocyte adhesion molecule VLA-4. Cloned cDNA
XX sequences of mouse 21.6 VH and VL (see AAQ99889) regions are linked to
XX human constant regions in the construction of a humanized antibody
XX against VLA-4. The 5' and 3' ends of the mouse cDNAs are modified using
XX PCR primers (See AAQ99895-98) and then subcloned into mammalian cell
XX expression vectors containing human kappa or gamma-1 constant regions. In
XX the humanized heavy chain, amino acids H27, H28, H29, H30, H44 and H71 in
XX the human HC VR framework are replaced by the amino acid present in the
XX equivalent position of the mouse 21.6 Ig H chain. Plasmids encoding the
XX chimeric antibodies are transfected into COS cells. The humanized
XX antibodies can be used to inhibit adhesion of a leukocyte to an
XX endothelial cell and to treat inflammatory diseases such as multiple
XX sclerosis. They can also be used in the treatment of stroke, cerebral
XX traumas, meningitis or encephalitis. The antibodies can also be used for
XX detecting VLA-4, for affinity purification or for generating anti-
XX idiotype antibodies
XX
XX Sequence 140 AA;
XX
XX Query Match      77.2%; Score 71; DB 2; Length 140;
XX Best Local Similarity 81.2%; Pred. No. 0.0022;
XX Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX Qy      1 KIDPANGKTKYDPFIQ 16
XX Db      69 RIDPANGYTKYDPKFK 84
XX
XX RESULT 30
XX AAW22410
XX ID AAW22410 standard; protein; 140 AA.
XX
XX AC AAW22410;
XX
XX DT 08-DEC-1997 (first entry)

```

---

```

XX Alpha-4 integrin mouse MAb 21.6 VH region.
XX
XX Alpha-4 integrin; humanised antibody; monoclonal antibody 21.6; asthma;
XX atherosclerosis; AIDS; dementia; diabetes; tumour; metastasis;
XX inflammatory bowel disease; rheumatoid arthritis; transplant rejection;
XX graft versus host disease; nephritis; atopic dermatitis; psoriasis;
XX myocardial ischaemia; acute leukocyte mediated lung injury; therapy.
XX
XX Mus musculus.
XX
XX Key      Location/Qualifiers
XX Peptide  1. .19
XX Region   20. .49
XX          /label= Leader
XX          /label= FR1
XX          /note= "framework region 1"
XX Region   50. .54
XX          /label= CDR1
XX          /note= "complementarity determining region 1"
XX Region   55. .68
XX          /label= FR2
XX          /note= "framework region 2"
XX Region   69. .85
XX          /label= CDR2
XX          /note= "complementarity determining region 2"
XX Region   86. .117
XX          /label= FR3
XX          /note= "framework region 3"
XX Region   118. .131
XX          /label= CDR3
XX          /note= "complementarity determining region 3"
XX Region   132. .140
XX          /label= FR4
XX          /note= "framework region 4"
XX
XX WO9718838-A1.
XX
XX 29-MAY-1997.
XX
XX 21-NOV-1996; 96WO-US018807.
XX
XX 21-NOV-1995; 95US-00561521.
XX
XX (ATHE-) ATHENA NEUROSCIENCES INC.
XX
XX Bendig MM, Leger OJ, Saldanha J, Jones ST, Yednock TA;
XX
XX WPI; 1997-297879/27.
XX N-PSDB; AAT74760.
XX
XX Uses of humanised alpha-4 integrin antibody - for treatment of asthma,
XX atherosclerosis, AIDS, dementia, etc.
XX
XX Claim 18; Page 69-70; 107pp; English.
XX
XX This polypeptide comprises the heavy chain variable region (VH) of mouse
XX anti-alpha-4 integrin monoclonal antibody 21.6. The complementarity
XX determining regions (CDRs) of the 21.6 VH can be incorporated into a
XX human 21/28'CL framework to produce a claimed humanised 21.6 VH (see
XX AAW22413) and a claimed humanised 21.6 antibody that is used in the
XX manufacture of a medicament for treating a disease selected from asthma,
XX atherosclerosis, AIDS, dementia, diabetes, inflammatory bowel disease,
XX rheumatoid arthritis, transplant rejection, graft versus host disease,
XX tumour metastasis, nephritis, atopic dermatitis, psoriasis, myocardial
XX ischaemia, and acute leukocyte mediated lung injury. The antibody may
XX also be used in the affinity purification of alpha-4 integrin for use as
XX a vaccine or an immunogen. It is also useful for generating idiotypic
XX antibodies. The humanised antibodies of the invention have a half-life in
XX the human circulation essentially equivalent to that of naturally
XX occurring human antibodies
XX
XX Sequence 140 AA;

```



Query Match 77.2%; Score 71; DB 2; Length 140;  
Best Local Similarity 81.2%; Pred. No. 0.0022;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
Db :||||| ||||| ||  
69 RIDPANGTKYDKPKFQ 84

RESULT 31  
AAR81333  
ID AAR81333 standard; protein; 142 AA.  
XX  
AC AAR81333;  
DT 23-MAR-1996 (first entry)  
XX  
DE Human VLA-4 reshaped antibody 21.6 light heavy variable region.  
XX  
KW Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;  
KW antibody engineering.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..19  
FT /note= "signal peptide"  
FT Region 20..49  
FT /note= "framework region 1"  
FT Region 50..54  
FT /note= "complementarity determining region 1"  
FT Region 55..68  
FT /note= "framework region 2"  
FT Region 69..85  
FT /note= "complementarity determining region 2"  
FT Region 86..117  
FT /note= "framework region 3"  
FT Region 118..131  
FT /note= "complementarity determining region 3"  
FT Region 132..142  
FT /note= "framework region 4"  
XX  
PN W09519790-A1.  
XX  
XX 27-JUL-1995.  
XX  
XX 25-JAN-1995; 95WO-US001219.  
XX  
XX 25-JAN-1994; 94US-00186269.  
XX  
XX (ATHE-) ATHENA NEUROSCIENCES INC.  
XX  
XX Bendig MM, Leger OJ, Saldanha J, Jones TS;  
XX WPI; 1995-269276/35.  
XX N-PSDB; AAQ99894.  
XX  
XX New humanised antibodies against VLA-4 - used for inhibiting leukocyte  
XX adhesion to endothelial cells, partic. for treating inflammatory disease.  
XX  
XX Disclosure; Fig 11; 105pp; English.  
XX  
XX The sequence represents the human reshaped antibody 21.6 heavy chain  
XX variable region against leukocyte adhesion molecule VLA-4. Cloned cDNA  
XX sequences of mouse 21.6 VH (AAQ99892) and VL (AAQ99889) regions are  
XX linked to human constant regions in the construction of a humanized  
XX antibody against VLA-4. The 5' and 3' ends of the mouse cDNAs are  
XX modified using PCR primers (AAQ99895-98) and then subcloned into  
XX mammalian cell expression vectors containing human kappa or gamma-1  
XX constant regions. In the humanized heavy chain, amino acids H27, H28,  
XX H29, H30, H44 and H71 in the human HC VR framework are replaced by the  
XX amino acid present in the equivalent position of the mouse 21.6 Ig H

CC chain. Plasmids encoding the chimeric antibodies are transfected into COS  
CC cells. The humanized antibodies can be used to inhibit adhesion of a  
CC leukocyte to an endothelial cell and to treat inflammatory diseases such  
CC as multiple sclerosis. They can also be used in the treatment of stroke,  
CC cerebral traumas, meningitis or encephalitis. The antibodies can also be  
CC used for detecting VLA-4, for affinity purification or for generating  
CC anti-idiotype antibodies  
XX  
SQ Sequence 142 AA;  
Query Match 77.2%; Score 71; DB 2; Length 142;  
Best Local Similarity 81.2%; Pred. No. 0.0022;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
Db :||||| ||||| ||  
69 RIDPANGTKYDKPKFQ 84

RESULT 32  
AAW22428  
ID AAW22428 standard; protein; 142 AA.  
XX  
AC AAW22428;  
XX  
DT 09-DEC-1997 (first entry)  
XX  
DE Humanised alpha-4 integrin antibody 21.6 VL version Ha.  
XX  
KW Alpha-4 integrin; humanised antibody; monoclonal antibody 21.6; asthma;  
KW atherosclerosis; AIDS; dementia; diabetes; tumour; metastasis;  
KW inflammatory bowel disease; rheumatoid arthritis; transplant rejection;  
KW graft versus host disease; nephritis; atopic dermatitis; psoriasis;  
KW myocardial ischaemia; acute leukocyte mediated lung injury; therapy.  
XX  
OS Mus; musculus.  
OS Homo; sapiens.  
OS Synthetic.  
OS Chimeric.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..19  
FT /label= Leader  
FT Protein 20..142  
FT /label= Mat\_protein  
FT /note= "VH version Ha (Claim 25)"  
FT Region 20..49  
FT /label= FR1  
FT /note= "21/28' CL framework region 1"  
FT Region 50..55  
FT /label= CDR1  
FT /note= "21.6 complementarity determining region 1"  
FT Region 55..67  
FT /label= FR2  
FT /note= "21/28' CL framework region 2"  
FT Region 68..85  
FT /label= CDR2  
FT /note= "21.6 complementarity determining region 2"  
FT Region 86..117  
FT /label= FR3  
FT /note= "21/28' CL framework region 3"  
FT Region 118..131  
FT /label= CDR3  
FT /note= "21.6 complementarity determining region 3"  
FT Region 132..142  
FT /label= FR4  
FT /note= "21/28' CL framework region 4"  
XX  
XX W09718838-A1.  
XX  
XX 29-MAY-1997.  
XX  
XX 21-NOV-1996; 96WO-US018807.

```

XX 21-NOV-1995; 95US-00561521.
XX (ATHE-) ATHENA NEUROSCIENCES INC.
XX Bendig MM, Leger OJ, Saldanha J, Jones ST, Yednock TA;
XX WPI; 1997-297879/27.
XX N-PSDB; AAT74789.
XX Uses of humanised alpha-4 integrin antibody - for treatment of asthma,
XX atherosclerosis, AIDS, dementia, etc.
XX Example 6; Fig 11; 107pp; English.
XX This polypeptide, designated Ha, comprises the heavy chain variable
XX region (VH) of a humanised alpha-4 integrin antibody 21.6 (see also
XX AAW22413). It is composed of complementarity determining regions from the
XX VH region (see AAW22410) of mouse alpha-4 integrin monoclonal antibody
XX 21.6 and a modified human 21/28'CL framework. It can be expressed in
XX mammalian host cells following PCR amplification and mutagenesis of
XX appropriate mouse and human DNA sequences. The humanised 21.6 VH and a
XX humanised 21.6 VL (see AAW22412) can be used to produce a claimed
XX humanised 21.6 antibody that is useful in the manufacture of a medicament
XX for treating asthma, atherosclerosis, AIDS, dementia, diabetes,
XX inflammatory bowel disease, rheumatoid arthritis, transplant rejection,
XX graft versus host disease, tumour metastasis, nephritis, atopic
XX dermatitis, psoriasis, myocardial ischaemia, and acute leukocyte mediated
XX lung injury. The humanised antibody has a half-life in the human
XX circulation essentially equivalent to that of naturally occurring human
XX antibodies
XX
XX Query Match 77.2%; Score 71; DB 2; Length 142;
XX Best Local Similarity 81.2%; Pred. No. 0.0022;
XX Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 KIDPANGKTKYDPIFQ 16
XX :|||||
XX Db 69 RIDPANGTKYDKPKFQ 84
XX
XX RESULT 33
XX ADR40432
XX ID ADR40432 standard; protein; 142 AA.
XX AC ADR40432;
XX
XX DT 04-NOV-2004 (first entry)
XX
XX DE Reshaped human (RH) 21.6 antibody Vh region version a sequence.
XX
XX KW Sulphonylamine; remyelination; demyelinating disease;
XX lymphocyte infiltration; paralysis; neuroprotective; muscular;
XX cyostatic; antiinflammatory; vulneryary;
XX KW selective adhesion molecule inhibitor; SAMI; antibody; 21.6; human.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT Peptide 1..19
XX FT /note= "leader sequence"
XX FT Region 20..49
XX FT /note= "FR1"
XX FT Region 50..54
XX FT /note= "CDR1"
XX FT Region 55..68
XX FT /note= "FR2"
XX FT Region 69..85
XX FT /note= "CDR2"
XX FT Region 86..117
XX FT /note= "FR3"

```

```

XX Region 118..131
XX FT /note= "CDR3"
XX FT 132..142
XX FT /note= "FR4"
XX
XX PN WO2004066931-A2.
XX
XX PD 12-AUG-2004.
XX
XX PF 26-JAN-2004; 2004WO-US002028.
XX
XX PR 24-JAN-2003; 2003US-0442171P.
XX PR 05-SEP-2003; 2003US-0500316P.
XX
XX PA (ELAN-) ELAN PHARM INC.
XX
XX PI Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;
XX PI Messersmith E, Freedman S, Yednock T;
XX
XX DR WPI; 2004-593990/57.
XX DR N-PSDB; ADR40431.
XX
XX PT Use of sulfonylamine compounds as selective adhesion molecule inhibitors
XX PT to treat demyelinating disease e.g. multiple sclerosis, by inhibiting
XX PT lymphocyte infiltration and promoting remyelination to treat paralysis.
XX
XX Example 3; Fig 16A-B; 573pp; English.
XX
XX CC The invention relates to sulphonylamine compounds (A) of specified
XX CC formulae. The compounds are used to prepare a medicament to promote
XX CC remyelination of nerve cells in a mammal for treatment of a demyelinating
XX CC disease and to inhibit lymphocyte infiltration of immune cells in the
XX CC spinal cord to promote remyelination of nerve cells in the spinal cord,
XX CC thus treating paralysis. The method of inhibiting lymphocyte infiltration
XX CC of immune cells in the spinal cord to promote remyelination of nerve
XX CC cells in the spinal cord and thus treating paralysis further comprises co
XX CC -administering an immunosuppressant such as adrenocorticotrophic hormone,
XX CC a corticosteroid (e.g. prednisone, methylprednisolone, dexamethasone
XX CC cortisol, cortisone, fludrocortisone, prednisolone, alpha-
XX CC methylprednisolone, triamcinolone or betamethasone) or an interferon
XX CC (e.g. interferon beta-1b or interferon beta-1a). (A) is used to prepare a
XX CC medicament for treating paralysis, a congenital metabolic disorder, a
XX CC neuropathy with abnormal myelination, drug induced demyelination,
XX CC radiation induced demyelination, hereditary demyelinating condition,
XX CC prion induced demyelinating condition, encephalitis induced demyelination
XX CC or spinal cord injury. The present sequence represents the first version
XX CC (a) of a reshaped human 21.6 antibody heavy chain variable region.
XX
XX SQ Sequence 142 AA;
XX
XX Query Match 77.2%; Score 71; DB 8; Length 142;
XX Best Local Similarity 81.2%; Pred. No. 0.0022;
XX Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 KIDPANGKTKYDPIFQ 16
XX :|||||
XX Db 69 RIDPANGTKYDKPKFQ 84
XX
XX RESULT 34
XX ADR40350
XX ID ADR40350 standard; protein; 142 AA.
XX
XX AC ADR40350;
XX
XX DT 04-NOV-2004 (first entry)
XX
XX DE Reshaped human (RH) 21.6 antibody Vh region version a sequence.
XX
XX KW Remyelination; demyelinating disease; lymphocyte infiltration; paralysis;
XX neuroprotective; muscular; cyostatic; antiinflammatory; vulneryary;
XX KW selective adhesion molecule inhibitor; SAMI; antibody; 21.6; human.
XX
XX

```

```

OS XX Homo sapiens.
PH Key Location/Qualifiers
FT Peptide 1..19
FT Region /note= "leader sequence"
FT FT 20..49
FT FT /note= "FR1"
FT FT 50..54
FT FT /note= "CDR1"
FT FT 55..68
FT FT /note= "FR2"
FT FT 69..85
FT FT /note= "CDR2"
FT FT 86..117
FT FT /note= "FR3"
FT FT 118..131
FT FT /note= "CDR3"
FT FT 132..142
FT FT /note= "FR4"
XX WO200406932-A2.
XX 12-AUG-2004.
XX 26-JAN-2004; 2004WO-US002039.
XX 24-JAN-2003; 2003US-0442171P.
XX 05-SEP-2003; 2003US-0500316P.
XX (ELAN-) ELAN PHARM INC.
XX Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;
XX Messersmith E, Freedman S, Yednock T;
XX WPI; 2004-593991/57.
XX N-PSDB; ADR40349.
XX Use of remyelinating agent such as natalizumab for preparation of
XX medicament for treating demyelinating diseases and paralysis by promoting
XX remyelination of nerve cells in mammal.
XX Example 3; Fig 16A-B; 495pp; English.
XX The invention relates to remyelinating agents (A) that are used to
XX prepare a medicament to promote remyelination of nerve cells in a mammal
XX for treatment of a demyelinating disease and to inhibit lymphocyte
XX infiltration of immune cells in the spinal cord to promote remyelination
XX of nerve cells in the spinal cord, thus treating paralysis. The method of
XX inhibiting lymphocyte infiltration of immune cells in the spinal cord to
XX promote remyelination of nerve cells in the spinal cord and thus treating
XX paralysis further comprises co-administering an immunosuppressant such as
XX adrenocorticotrophic hormone, a corticosteroid (e.g. prednisone,
XX methylprednisolone, dexamethasone, cortisone, fludrocortisone,
XX prednisolone, alpha-methylprednisolone, triamcinolone or betamethasone)
XX or an interferon (e.g. interferon beta-1b or interferon beta-1a). (A) is
XX used to prepare a medicament for treating paralysis, a congenital
XX metabolic disorder, a neuropathy with abnormal myelination, drug induced
XX demyelination, radiation induced demyelination, hereditary demyelinating
XX condition, prion induced demyelinating condition, encephalitis induced
XX demyelination or spinal cord injury. The present sequence represents the
XX first version (a) of a reshaped human 21.6 antibody heavy chain variable
XX region.
XX SQ Sequence 142 AA;
XX Query Match 77.2%; Score 71; DB 8; Length 142;
XX Best Local Similarity 81.2%; Pred. No. 0.0022;
XX Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Oy 1 KIDPANGTKYDPIFQ 16
Db :|||||
69 RIDPANGTKYDKPFQ 84

```

---

```

RESULT 35
ADR40419
ID ADR40419 standard; protein; 144 AA.
XX AC ADR40419;
XX DT 04-NOV-2004 (first entry)
XX DE Mouse 21.6 antibody Vh region sequence.
XX KW Sulphonylamine; remyelination; demyelinating disease;
XX KW lymphocyte infiltration; paralysis; neuroprotective; muscular;
XX KW cytostatic; antiinflammatory; vulnery;
XX KW selective adhesion molecule inhibitor; SAMI; antibody; 21.6; mouse; gene;
XX KW ds.
XX OS Mus sp.
XX PH Key Location/Qualifiers
FT Peptide 1..19
FT Region /note= "leader sequence"
FT FT 20..49
FT FT /note= "FR1"
FT FT 50..54
FT FT /note= "CDR1"
FT FT 55..68
FT FT /note= "FR2"
FT FT 69..85
FT FT /note= "CDR2"
FT FT 86..117
FT FT /note= "FR3"
FT FT 118..131
FT FT /note= "CDR3"
FT FT 132..144
FT FT /note= "FR4"
FT FT Misc-difference 143
FT FT /note= "encoded by GCC"
FT FT Misc-difference 144
FT FT /note= "encoded by AAA"
XX WO2004066931-A2.
XX 12-AUG-2004.
XX 26-JAN-2004; 2004WO-US002028.
XX 24-JAN-2003; 2003US-0442171P.
XX 05-SEP-2003; 2003US-0500316P.
XX (ELAN-) ELAN PHARM INC.
XX Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;
XX Messersmith E, Freedman S, Yednock T;
XX WPI; 2004-593990/57.
XX N-PSDB; ADR40418.
XX Use of sulfonylamine compounds as selective adhesion molecule inhibitors
XX to treat demyelinating disease e.g. multiple sclerosis, by inhibiting
XX lymphocyte infiltration and promoting remyelination to treat paralysis.
XX Example 3; Fig 12A-B; 573pp; English.
XX The invention relates to sulphonylamine compounds (A) of specified
XX formulae. The compounds are used to prepare a medicament to promote
XX remyelination of nerve cells in a mammal for treatment of a demyelinating
XX disease and to inhibit lymphocyte infiltration of immune cells in the
XX spinal cord to promote remyelination of nerve cells in the spinal cord,
XX thus treating paralysis. The method of inhibiting lymphocyte infiltration
XX of immune cells in the spinal cord to promote remyelination of nerve
XX cells in the spinal cord and thus treating paralysis further comprises co
XX -administering an immunosuppressant such as adrenocorticotrophic hormone,

```

CC a corticosteroid (e.g. prednisone, methylprednisolone, dexamethasone  
 CC cortisol, cortisone, fluocortisone, prednisolone, alpha-  
 CC methylprednisolone, triamcinolone or betamethasone) or an interferon  
 CC (e.g. interferon beta-1b or interferon beta-1a). (A) is used to prepare a  
 CC medicament for treating paralysis, a congenital metabolic disorder, a  
 CC neuropathy with abnormal myelination, drug induced demyelination,  
 CC radiation induced demyelination, hereditary demyelinating condition,  
 CC prion induced demyelinating condition, encephalitis induced demyelination  
 CC or spinal cord injury. The present sequence represents a mouse 21.6  
 CC antibody heavy chain variable region (Vh) sequence.  
 XX  
 SQ Sequence 144 AA;

Query Match 77.2%; Score 71; DB 8; Length 144;  
 Best Local Similarity 81.2%; Pred. No. 0.0023;  
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16  
 :||||| ||||| ||  
 Db 69 RIDPANGYTKYDPKFQ 84

RESULT 37

ADR40337  
 ID ADR40337 standard; protein; 144 AA.

XX ADR40337;

DT 04-NOV-2004 (first entry)

XX Mouse 21.6 antibody Vh region sequence.

XX Remyelination; demyelinating disease; lymphocyte infiltration; paralysis;  
 KW neuroprotective; muscular; cytostatic; antiinflammatory; vulneryary;  
 KW selective adhesion molecule inhibitor; SAMI; antibody; 21.6; mouse; gene;  
 KW ds.

XX Mus sp.

XX Key Location/Qualifiers  
 FH Peptide 1..19  
 FT /note= "leader sequence"  
 FT Region 20..49  
 FT /note= "FR1"  
 FT Region 50..54  
 FT /note= "CDR1"  
 FT Region 55..68  
 FT /note= "FR2"  
 FT Region 69..85  
 FT /note= "CDR2"  
 FT Region 86..117  
 FT /note= "FR3"  
 FT Region 118..131  
 FT /note= "CDR3"  
 FT Region 132..144  
 FT /note= "FR4"  
 FT Misc-difference 143  
 FT /note= "encoded by GCC"  
 FT Misc-difference 144  
 FT /note= "encoded by AAA"  
 FT  
 XX WO2004066932-A2.

XX 12-AUG-2004.

XX 26-JAN-2004; 2004WO-US002039.  
 XX 24-JAN-2003; 2003US-0442171P.  
 XX 05-SEP-2003; 2003US-0500316P.

XX (ELAN-) ELAN PHARM INC.

XX Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;  
 PI

PI Messersmith E, Freedman S, Yednock T;

XX WPI; 2004-593991/57.

DR N-PSDB; ADR40336.

XX Use of remyelinating agent such as natalizumab for preparation of  
 PT medicament for treating demyelinating diseases and paralysis by promoting  
 PT remyelination of nerve cells in mammal.

XX Example 3; Fig 12A-B; 495pp; English.

XX The invention relates to remyelinating agents (A) that are used to  
 CC prepare a medicament to promote remyelination of nerve cells in a mammal  
 CC for treatment of a demyelinating disease and to inhibit lymphocyte  
 CC infiltration of immune cells in the spinal cord to promote remyelination  
 CC of nerve cells in the spinal cord, thus treating paralysis. The method of  
 CC inhibiting lymphocyte infiltration of immune cells in the spinal cord to  
 CC promote remyelination of nerve cells in the spinal cord and thus treating  
 CC paralysis further comprises co-administering an immunosuppressant such as  
 CC adrenocorticotrophic hormone, a corticosteroid (e.g. prednisone,  
 CC methylprednisolone, dexamethasone cortisol, cortisone, fludrocortisone,  
 CC prednisolone, alpha-methylprednisolone, triamcinolone or betamethasone)  
 CC or an interferon (e.g. interferon beta-1b or interferon beta-1a). (A) is  
 CC used to prepare a medicament for treating paralysis, a congenital  
 CC metabolic disorder, a neuropathy with abnormal myelination, drug induced  
 CC demyelination, radiation induced demyelination, hereditary demyelinating  
 CC condition, prion induced demyelinating condition, encephalitis induced  
 CC demyelination or spinal cord injury. The present sequence represents a  
 CC mouse 21.6 antibody heavy chain variable region (Vh) sequence.

XX Sequence 144 AA;

Query Match 77.2%; Score 71; DB 8; Length 144;

Best Local Similarity 81.2%; Pred. No. 0.0023;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16

:||||| ||||| ||  
 Db 69 RIDPANGYTKYDPKFQ 84

RESULT 37

ABR61875

ID ABR61875 standard; peptide; 17 AA.

XX ABR61875;

DT 12-SEP-2003 (first entry)

XX Mouse MAb 11K2 heavy chain variable region CDR2 fragment.

XX Antibody; beta-chemokine; monocyte chemotactic protein; MCP-1; MCP-2;  
 KW MCP-3; antiinflammatory; nephrotrophic; dermatologic; hepatotrophic;  
 KW neuroprotective; antiatherosclerotic; antirheumatic; antiarthritic;  
 KW cytostatic; anticonvulsant; antidiabetic; ophthalmological; vasotropic;  
 KW immunomodulator; cerebroprotective; antiulcer; antiparkinsonian; MAb;  
 KW monoclonal antibody; 11K2; complementarity determining region; CDR.

XX Mus musculus.

XX WO2003048083-A2.

XX 12-JUN-2003.

XX 27-NOV-2002; 2002WO-US038229.

XX 30-NOV-2001; 2001US-0343391P.

XX 24-MAY-2002; 2002US-0383277P.

XX 01-AUG-2002; 2002US-0400469P.

XX (BIOJ ) BIOGEN INC.

XX De Fougereolles AR, Kotelianski VE, Garber E, Reid C;

XX WPI; 2003-532819/50.  
 XX Novel antibody or its antigen binding fragment that specifically binds to  
 PT monocytic chemotactic protein-2 and other beta-chemokine, useful for  
 PT detecting presence of beta-chemokine in sample, and blocking chemotaxis.  
 XX  
 PS Claim 51; Page 70; 100pp; English.  
 XX  
 CC The invention relates to an antibody or its antigen binding fragment (AF)  
 CC that specifically binds to a plurality of beta-chemokines, particularly  
 CC monocytic chemotactic proteins MCP-1, MCP-2 and MCP-3. The antibodies are  
 CC useful in an immunoassay method for detecting the presence of a beta-  
 CC chemokine in a sample, and for blocking chemotaxis. The antibodies or  
 CC their Afs are useful for treating a subject suffering from a disorder  
 CC chosen from glomerulonephritis, scleroderma, multiple sclerosis, lupus  
 CC nephritis, cirrhosis, atherosclerosis, inflammatory bowel disease or  
 CC rheumatoid arthritis. The antibodies are also useful for treating  
 CC oncogenic disease or cancer, Crohn's disease, diabetic nephropathy,  
 CC cachexia, stroke, ulcerative colitis, coronary restenosis, Huntington's  
 CC disease and Parkinson's disease. Sequences ABR618/4-76 represent murine  
 CC monoclonal antibody (MAb) 11K2 heavy chain variable region CDRI-3  
 CC (complementarity determining region) fragments  
 XX  
 SQ Sequence 17 AA;  
 Query Match 75.0%; Score 69; DB 7; Length 17;  
 Best Local Similarity 75.0%; Pred. No. 0.0005;  
 Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KIDPANGTKYVDPIFQ 16  
 :|||||:|:|:|  
 Db 1 RIDPANGTKFDPKRFQ 16  
 RESULT 38  
 ADQ31264  
 ID ADQ31264 standard; peptide; 17 AA.  
 XX  
 AC ADQ31264;  
 XX  
 DT 09-SEP-2004 (first entry)  
 XX  
 DE CDR2 peptide of murine 11K2 heavy chain variable domain antibody Seq 30.  
 XX  
 KW CDR; 11K2; monocytic chemotactic protein; beta-chemokine family;  
 KW glomerulonephritis; scleroderma; cirrhosis; multiple sclerosis;  
 KW lupus nephritis; atherosclerosis; inflammatory bowel disease;  
 KW rheumatoid arthritis; inflammatory disease; fibrotic disorder; cancer;  
 KW immunopathological disorder; antiarteriosclerotic; antiarthritic;  
 KW antiinflammatory; antirheumatic; cytostatic; dermatological;  
 KW hepatotropic; immunomodulator; nephrotropic; neuroprotective; mouse; MCP;  
 KW murine.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO2004050836-A2.  
 XX  
 PD 17-JUN-2004.  
 XX  
 PP 25-NOV-2003; 2003WO-US037834.  
 XX  
 PR 27-NOV-2002; 2002US-0430007P.  
 XX  
 PA (BIOG-) BIOGEN IDEC MA INC.  
 XX  
 PI De Fougereolles AR, Kotelianski VE, Garber E, Reid C, Saldanha JW;  
 PI Van Vlijmen H;  
 XX  
 DR WPI; 2004-461110/43.  
 XX  
 XX New antibodies against monocytic chemotactic proteins (MCP), useful for  
 PT treating or preventing disorders associated with detrimental MCP

PT activity, e.g. glomerulonephritis, scleroderma, multiple sclerosis, or  
 XX atherosclerosis.  
 XX  
 PS Claim 8; SEQ ID NO 30; 200pp; English.  
 XX  
 CC This invention relates to an antibody for treating or preventing  
 CC disorders associated with detrimental monocytic chemotactic protein (MCP)  
 CC activity. Specifically, it refers to humanised antibodies that bind to  
 CC members of the beta-chemokine family (of which MCP-1, MCP-2 and MCP-3  
 CC belong) and in particular antibodies that have been modelled on, and  
 CC modified from, the variable complementarity determining regions (CDRs) of  
 CC the murine 11K2 and 1A1 immunoglobulin sequences. The present invention  
 CC describes using these antibodies to treat or prevent diseases and  
 CC disorders including glomerulonephritis, scleroderma, cirrhosis, multiple  
 CC sclerosis, lupus nephritis, atherosclerosis, inflammatory bowel diseases,  
 CC rheumatoid arthritis, inflammatory diseases, fibrotic disorders, cancer  
 CC and immunopathological disorders. Accordingly, they can be used in the  
 CC development of pharmaceutical compositions that exhibit  
 CC antiarteriosclerotic, antiarthritic, antiinflammatory, antirheumatic,  
 CC cytostatic, dermatological, hepatotropic, immunomodulator, nephrotropic  
 CC and neuroprotective activities. This peptide sequence is a CDR region of  
 CC the murine 11K1 heavy chain antibody protein of the invention.  
 XX  
 SQ Sequence 17 AA;  
 Query Match 75.0%; Score 69; DB 8; Length 17;  
 Best Local Similarity 75.0%; Pred. No. 0.0005;  
 Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KIDPANGTKYVDPIFQ 16  
 :|||||:|:|:|  
 Db 1 RIDPANGTKFDPKRFQ 16  
 RESULT 39  
 ABR61872  
 ID ABR61872 standard; protein; 117 AA.  
 XX  
 AC ABR61872;  
 XX  
 DT 12-SEP-2003 (first entry)  
 XX  
 DE Mouse MAb 11K2 heavy chain variable region.  
 XX  
 KW Antibody; beta-chemokine; monocytic chemotactic protein; MCP-1; MCP-2;  
 KW MCP-3; antiinflammatory; nephrotropic; dermatological; hepatotropic;  
 KW neuroprotective; antiarteriosclerotic; antirheumatic; antiarthritic;  
 KW cytostatic; anticonvulsant; antidiabetic; ophthalmological; vasotropic;  
 KW immunomodulator; cerebroprotective; antiulcer; antiparkinsonian; MAb;  
 KW monoclonal antibody; 11K2.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO2003048083-A2.  
 XX  
 PD 12-JUN-2003.  
 XX  
 PP 27-NOV-2002; 2002WO-US038229.  
 XX  
 PR 30-NOV-2001; 2001US-0343391P.  
 PR 24-MAY-2002; 2002US-0383277P.  
 PR 01-AUG-2002; 2002US-0400469P.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI De Fougereolles AR, Kotelianski VE, Garber E, Reid C;  
 XX  
 DR WPI; 2003-532819/50.  
 DR N-PSDE; ACC84737.  
 XX  
 XX Novel antibody or its antigen binding fragment that specifically binds to  
 PT monocytic chemotactic protein-2 and other beta-chemokine, useful for  
 PT detecting presence of beta-chemokine in sample, and blocking chemotaxis.

XX Claim 42; Page 69; 100pp; English.  
XX The invention relates to an antibody or its antigen binding fragment (AF)  
CC that specifically binds to a plurality of beta-chemokines, particularly  
CC monocyte chemotactic proteins MCP-1, MCP-2 and MCP-3. The antibodies are  
CC useful in an immunoassay method for detecting the presence of a beta-  
CC chemokine in a sample, and for blocking chemotaxis. The antibodies or  
CC their Afs are useful for treating a subject suffering from a disorder  
CC chosen from glomerulonephritis, scleroderma, multiple sclerosis, lupus  
CC nephritis, cirrhosis, atherosclerosis, inflammatory bowel disease or  
CC rheumatoid arthritis. The antibodies are also useful for treating  
CC oncogenic disease or cancer, Crohn's disease, diabetic nephropathy,  
CC cachexia, stroke, ulcerative colitis, coronary restenosis, Huntington's  
CC disease and Parkinson's disease. The present sequence represents a murine  
CC monoclonal antibody (MAb) 11K2 heavy chain variable region  
XX  
SQ Sequence 117 AA;  
  
Query Match 75.0%; Score 69; DB 7; Length 117;  
Best Local Similarity 75.0%; Pred. No. 0.0038;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 KIDPANGTKYDPIFQ 16  
Db :||||| ||:| |  
50 RIDPANGNTKFDPKFQ 65  
  
RESULT 40  
ADQ31281  
ID ADQ31281 standard; protein; 117 AA.  
XX  
AC ADQ31281;  
XX  
DT 09-SEP-2004 (first entry)  
XX  
DE Humanised murine 11K2 heavy chain variable domain antibody (v1) SeqID 47.  
XX  
KW 11K2; monocyte chemotactic protein; beta-chemokine family;  
KW glomerulonephritis; scleroderma; cirrhosis; multiple sclerosis;  
KW lupus nephritis; atherosclerosis; inflammatory bowel disease;  
KW rheumatoid arthritis; inflammatory disease; fibrotic disorder; cancer;  
KW immunopathological disorder; antiarteriosclerotic; antiarthritic;  
KW antiinflammatory; antirheumatic; cytostatic; dermatological;  
KW hepatotropic; immunomodulator; nephrotropic; neuroprotective; mouse; MCP;  
KW murine; humanised antibody.  
XX  
OS Mus musculus.  
OS Synthetic.  
XX  
XX WO2004050836-A2.  
XX 17-JUN-2004.  
XX  
XX 25-NOV-2003; 2003WO-US037834.  
XX  
XX 27-NOV-2002; 2002US-0430007P.  
XX  
XX (BIOGEN) BIOGEN IDEC MA INC.  
XX  
XX De Fougereolles AR, Kotellianski VE, Garber E, Reid C, Saldanha JW;  
XX Van Vlijmen H;  
XX WPI; 2004-461110/43.  
XX  
XX New antibodies against monocyte chemotactic proteins (MCP), useful for  
XX treating or preventing disorders associated with detrimental MCP  
XX activity, e.g. glomerulonephritis, scleroderma, multiple sclerosis, or  
XX atherosclerosis.  
XX  
XX Claim 54; SEQ ID NO 47; 200pp; English.  
XX  
XX This invention relates to an antibody for treating or preventing

disorders associated with detrimental monocyte chemotactic protein (MCP) activity. Specifically, it refers to humanised antibodies that bind to members of the beta-chemokine family (of which MCP-1, MCP-2 and MCP-3 belong) and in particular antibodies that have been modelled on, and modified from, the variable complementarity determining regions (CDRs) of the murine 11K2 and 1A1 immunoglobulin sequences. The present invention describes using these antibodies to treat or prevent diseases and disorders including glomerulonephritis, scleroderma, cirrhosis, multiple sclerosis, lupus nephritis, atherosclerosis, inflammatory bowel disease, rheumatoid arthritis, inflammatory diseases, fibrotic disorders, cancer and immunopathological disorders. Accordingly, they can be used in the development of pharmaceutical compositions that exhibit antiarteriosclerotic, antiarthritic, antiinflammatory, antirheumatic, cytostatic, dermatological, hepatotropic, immunomodulator, nephrotropic and neuroprotective activities. This polypeptide sequence is a humanised murine 11K2 antibody protein (version 1) of the invention.

Query Match 75.0%; Score 69; DB 8; Length 117;  
Best Local Similarity 75.0%; Pred. No. 0.0038;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 KIDPANGTKYDPIFQ 16  
Db :||||| ||:| |  
50 RIDPANGNTKFDPKFQ 65  
  
Search completed: June 3, 2005, 15:56:44  
Job time : 158.606 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 15:49:21 ; Search time 39.6667 Seconds  
(without alignments)  
31.992 Million cell updates/sec

Title: US-10-089-452-22

Perfect score: 92

Sequence: 1 KIDPANGKTKYDPIFQA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	79.3	18	1	Sequence 14, Appl
2	73	79.3	113	1	Sequence 2, Appli
3	73	79.3	125	2	Sequence 44, Appl
4	73	79.3	125	5	Sequence 44, Appl
5	73	79.3	136	1	Sequence 2, Appli
6	71	77.2	123	2	Sequence 9, Appli
7	71	77.2	123	2	Sequence 11, Appl
8	71	77.2	123	5	Sequence 9, Appli
9	71	77.2	123	5	Sequence 11, Appl
10	71	77.2	140	2	Sequence 4, Appli
11	71	77.2	140	5	Sequence 4, Appli
12	71	77.2	142	2	Sequence 17, Appl
13	71	77.2	142	5	Sequence 17, Appl
14	70	76.1	109	4	Sequence 28, Appl
15	69	75.0	254	2	Sequence 4, Appli
16	69	75.0	254	2	Sequence 7, Appli
17	69	75.0	254	2	Sequence 10, Appl
18	69	75.0	254	2	Sequence 13, Appl
19	68	73.9	17	3	Sequence 47, Appl
20	68	73.9	17	3	Sequence 116, App
21	68	73.9	17	4	Sequence 47, Appl
22	68	73.9	117	3	Sequence 132, App
23	68	73.9	136	3	Sequence 29, Appl
24	68	73.9	136	3	Sequence 99, Appl
25	68	73.9	137	3	Sequence 31, Appl
26	68	73.9	137	4	Sequence 31, Appl
27	68	73.9	269	3	Sequence 109, App

US-08-024-253-14  
; Sequence 14, Application US/08024253  
; Patent No. 5785968  
; GENERAL INFORMATION:  
; APPLICANT: KIMACHI, Kazuhiko  
; APPLICANT: MAEDA, Hiroaki  
; APPLICANT: NISHIYAMA, Kiyoto  
; APPLICANT: TOKIYOSHI, Sachio  
; APPLICANT: TOHYA, Yukinobu  
; APPLICANT: MIKAMI, Takeshi  
; TITLE OF INVENTION: ANTI-PELVINE CALCIVIRUS RECOMBINANT  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WEGNER, CANTOR, MUELLER & PLAYER  
; STREET: 1233 20th Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20036-8218  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 19930301  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 79189/1992  
; FILING DATE: 28-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CANTOR, Herbert I.  
; REGISTRATION NUMBER: 24,392  
; REFERENCE/DOCKET NUMBER: P-500-23744  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-0400  
; TELEFAX: (202) 835-0605  
; TELEX: 440706 WEGBR  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-024-253-14

#### ALIGNMENTS

RESULT 1  
US-08-024-253-14  
; Sequence 14, Application US/08024253  
; Patent No. 5785968  
; GENERAL INFORMATION:  
; APPLICANT: KIMACHI, Kazuhiko  
; APPLICANT: MAEDA, Hiroaki  
; APPLICANT: NISHIYAMA, Kiyoto  
; APPLICANT: TOKIYOSHI, Sachio  
; APPLICANT: TOHYA, Yukinobu  
; APPLICANT: MIKAMI, Takeshi  
; TITLE OF INVENTION: ANTI-PELVINE CALCIVIRUS RECOMBINANT  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WEGNER, CANTOR, MUELLER & PLAYER  
; STREET: 1233 20th Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20036-8218  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 19930301  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 79189/1992  
; FILING DATE: 28-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CANTOR, Herbert I.  
; REGISTRATION NUMBER: 24,392  
; REFERENCE/DOCKET NUMBER: P-500-23744  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-0400  
; TELEFAX: (202) 835-0605  
; TELEX: 440706 WEGBR  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-024-253-14

Query Match 79.3% Score 73; DB 1; Length 18;

```

: TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
:
: TITLE OF INVENTION: Adhesion Molecule VLA-4
:
: NUMBER OF SEQUENCES: 45
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Townsend and Townsend Kourie and Crew
:
: STREET: One Market Plaza, Steuart Tower, Suite 2000
:
: CITY: San Francisco
:
: STATE: California
:
: COUNTRY: USA
:
: ZIP: 94105
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/561,521
:
: FILING DATE:
:
: CLASSIFICATION: 424
:
: PRIOR APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/186,269A
:
: FILING DATE: 25-JAN-1994
:
: ATTORNEY/AGENT INFORMATION:
:
: NAME: Smith, William L.
:
: REGISTRATION NUMBER: 30,223
:
: REFERENCE/DOCKET NUMBER: 15270-14
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: 415-543-9600
:
: TELEFAX: 415-543-5043
:
: INFORMATION FOR SEQ ID NO: 44:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 125 amino acids
:
: TYPE: amino acid
:
: STRANDEDNESS: single
:
: TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
: US-08-561-521-44
:
:
:
:
: Query Match 79.3% Score 73; DB 2; Length 125;
:
: Best Local Similarity 81.2%; Pred. No. 0.00016;
:
: Matches 13; Conservative 1; Mismatches 2; Indels
:
:
: QY 1 KIDPANGTKYDPIFQ 16
:
: :|||||
:
: Db 50 RIDPANGNTKYDPKFQ 65
:
:
:
: RESULT 4
:
: PCT-US95-01219-44
:
: Sequence 44, Application PC/TUS9501219
:
: GENERAL INFORMATION:
:
: APPLICANT: Bendig, Mary M.
:
: APPLICANT: Leger, Olivier J.
:
: APPLICANT: Saldanha, Jose
:
: APPLICANT: Jones, S. Tarran
:
: TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
:
: TITLE OF INVENTION: Adhesion Molecule VLA-4
:
: NUMBER OF SEQUENCES: 45
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Townsend and Townsend Kourie and Crew
:
: STREET: One Market Plaza, Steuart Tower, Suite 2000
:
: CITY: San Francisco
:
: STATE: California
:
: COUNTRY: USA
:
: ZIP: 94105
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: PCT/US95/01219
:
: FILING DATE: 25-JAN-1995
:
:

```



```
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; FILING DATE: 25-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William L.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 15270-14
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 125 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-01219-44

Query Match 79.3%; Score 73; DB 5; Length 125;
Best Local Similarity 81.2%; Pred. No. 0.00016;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIRFQ 16
Db 50 RIDPANGTKYDPRFQ 65

RESULT 5
US-08-024-253-2
; Sequence 2, Application US/08024253
; Patent No. 5785968
; GENERAL INFORMATION:
; APPLICANT: KIMACHI, Kazuhiko
; APPLICANT: MAEDA, Hiroaki
; APPLICANT: NISHIYAMA, Kiyoto
; APPLICANT: TOKIYOSHI, Sachio
; APPLICANT: TOHYA, Yukinobu
; APPLICANT: MIKAMI, Takeshi
; TITLE OF INVENTION: ANTI-PELVINE CALCIVIRUS RECOMBINANT
; TITLE OF INVENTION: ANTIBODY AND GENE FRAGMENT ENCODING THE SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WEGNER, CANTOR, MUELLER & PLAYER
; STREET: 1233 20th Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-8218
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19930301
; APPLICATION NUMBER: US/08/024,253
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 79189/1992
; FILING DATE: 28-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTOR, Herbert I.
; REGISTRATION NUMBER: 24,392
; REFERENCE/DOCKET NUMBER: P-500-23744
; TELEPHONE: (202) 887-0400
; TELEFAX: (202) 835-0605
; TELEX: 440706 WEGEBR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 136 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-024-253-2

Query Match 79.3%; Score 73; DB 1; Length 136;
Best Local Similarity 81.2%; Pred. No. 0.00017;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIRFQ 16
Db 69 RIDPANGTKYDPRFQ 84

RESULT 6
US-08-561-521-9
; Sequence 9, Application US/08561521
; Patent No. 5840299
; GENERAL INFORMATION:
; APPLICANT: Bendig, Mary M.
; APPLICANT: Leger, Olivier J.
; APPLICANT: Saldanha, Jose
; APPLICANT: Jones, S. Tarran
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
; TITLE OF INVENTION: Adhesion Molecule VLA-4
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/561,521
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,269A
; FILING DATE: 25-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William L.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 15270-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-561-521-9

Query Match 77.2%; Score 71; DB 2; Length 123;
Best Local Similarity 81.2%; Pred. No. 0.00033;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIRFQ 16
Db 50 RIDPANGTKYDPRFQ 65

RESULT 7
US-08-561-521-11
; Sequence 11, Application US/08561521
```

Patent No. 5840299  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.  
; APPLICANT: Saldanha, Jose  
; APPLICANT: Jones, S. Tarran  
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/561,521  
; FILING DATE: 25-JAN-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/186,269A  
; FILING DATE: 25-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William L.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15270-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 123 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-561-521-11

Query Match 77.2%; Score 71; DB 2; Length 123;  
Best Local Similarity 81.2%; Pred. No. 0.00033;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
Db 50 RIDPANGTKYDPKFK 65

RESULT 8  
PCT-US95-01219-9  
; Sequence 9, Application PC/TUS9501219  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.  
; APPLICANT: Saldanha, Jose  
; APPLICANT: Jones, S. Tarran  
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/01219  
; FILING DATE: 25-JAN-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/186,269  
; FILING DATE: 25-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William L.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15270-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 123 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-01219-9

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01219  
FILING DATE: 25-JAN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/186,269  
FILING DATE: 25-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William L.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 15270-14  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 123 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-01219-9

Query Match 77.2%; Score 71; DB 5; Length 123;  
Best Local Similarity 81.2%; Pred. No. 0.00033;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
Db 50 RIDPANGTKYDPKFK 65

RESULT 9  
PCT-US95-01219-11  
; Sequence 11, Application PC/TUS9501219  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.  
; APPLICANT: Saldanha, Jose  
; APPLICANT: Jones, S. Tarran  
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/01219  
; FILING DATE: 25-JAN-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/186,269  
; FILING DATE: 25-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William L.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15270-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 123 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-01219-11

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-01219-11
Query Match 77.2%; Score 71; DB 5; Length 123;
Best Local Similarity 81.2%; Pred. No. 0.00033;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 50 RIDPANGTKYDPKFQ 65

RESULT 10
US-08-561-521-4
; Sequence 4, Application US/08561521
; GENERAL INFORMATION:
; APPLICANT: Bendig, Mary M.
; APPLICANT: Leger, Olivier J.
; APPLICANT: Saldanha, Jose
; APPLICANT: Jones, S. Tarran
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01219
; FILING DATE: 25-JAN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/186,269
; FILING DATE: 25-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William L.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 15270-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 4:
; LENGTH: 140 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-01219-4
Query Match 77.2%; Score 71; DB 5; Length 140;
Best Local Similarity 81.2%; Pred. No. 0.00037;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 69 RIDPANGTKYDPKFQ 84

RESULT 12
US-08-561-521-17
; Sequence 17, Application US/08561521
; Patent No. 5840299
; GENERAL INFORMATION:
; APPLICANT: Bendig, Mary M.
; APPLICANT: Leger, Olivier J.
; APPLICANT: Saldanha, Jose
; APPLICANT: Jones, S. Tarran
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/561,521
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,269A
; FILING DATE: 25-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William L.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 15270-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-543-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 140 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-561-521-4
Query Match 77.2%; Score 71; DB 2; Length 140;
Best Local Similarity 81.2%; Pred. No. 0.00037;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 69 RIDPANGTKYDPKFQ 84

RESULT 11
```

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/561.521  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/186.269A  
; FILING DATE: 25-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William L.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15270-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 142 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-561-521-17

Query Match 77.2%; Score 71; DB 2; Length 142;  
Best Local Similarity 81.2%; Pred. No. 0.00038;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
Db 69 RIDPANGTKYDPKFQ 84

RESULT 13  
PCT-US95-01219-17  
; Sequence 17, Application PC/TUS9501219  
; GENERAL INFORMATION:  
; APPLICANT: Bendig, Mary M.  
; APPLICANT: Leger, Olivier J.  
; APPLICANT: Saldanha, Jose  
; APPLICANT: Jones, S. Tarran  
; TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
; TITLE OF INVENTION: Adhesion Molecule VLA-4  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/01219  
; FILING DATE: 25-JAN-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/186,269  
; FILING DATE: 25-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William L.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 15270-14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 142 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-01219-17

Query Match 77.2%; Score 71; DB 5; Length 142;  
Best Local Similarity 81.2%; Pred. No. 0.00038;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
Db 69 RIDPANGTKYDPKFQ 84

RESULT 14  
US-09-497-997C-28  
; Sequence 28, Application US/09497997C  
; Patent No. 6635248  
; GENERAL INFORMATION:  
; APPLICANT: TERNYNCK, THERESA  
; APPLICANT: AVRAMEAS, ALEXANDRE  
; APPLICANT: BUTTIN, GERARD  
; APPLICANT: AVRAMEAS, STRAITIS  
; APPLICANT: SARON, MARIE-FRANCOISE  
; APPLICANT: BLONDEL, BRUNO  
; APPLICANT: COUDERC, THERESA  
; APPLICANT: MICHELSON, SUSAN  
; APPLICANT: ZIPETO, DONATO  
; TITLE OF INVENTION: VECTOR DERIVED FROM ANTIBODIES FOR TRANSFERRING SUBSTANCES INTO C  
; FILE REFERENCE: 0660-0166-0XCONT  
; CURRENT APPLICATION NUMBER: US/09/497,997C  
; CURRENT FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: PCT/FR98/01740  
; PRIOR FILING DATE: 1998-08-04  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28  
; LENGTH: 109  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE  
; US-09-497-997C-28

Query Match 76.1%; Score 70; DB 4; Length 109;  
Best Local Similarity 81.2%; Pred. No. 0.00041;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
Db 41 RIDPANGTKYAPKFQ 56

RESULT 15  
US-08-792-824-4  
; Sequence 4, Application US/08792824  
; Patent No. 5932449  
; GENERAL INFORMATION:  
; APPLICANT: EMANUEL, PETER A.  
; APPLICANT: BURANS, JAMES P.  
; APPLICANT: VALDES, JAMES J.  
; APPLICANT: MOHYE, ELDEFRAWI E.  
; TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: U.S. Army Chemical and Biological Defense  
; ADDRESSEE: Command  
; STREET: Office of the Chief Counsel, Bldg E4435  
; CITY: Aberdeen Proving Ground  
; STATE: MD  
; COUNTRY: U.S.

```
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/792,824
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Biffoni, U. J.
; REGISTRATION NUMBER: 39,908
; REFERENCE/DOCKET NUMBER: DAM 431-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 410-671-1158
; TELEFAX: 410-671-2534
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-792-824-4
;
; Query Match 75.0%; Score 69; DB 2; Length 254;
; Best Local Similarity 75.0%; Pred. No. 0.0015;
; Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
;
QY 1 KIDPANGTKYDPIFQ 16
Db :|||||:|||||
72 RIDPANGNTEYDPKFQ 87

RESULT 16
US-08-792-824-7
; Sequence 7, Application US/08792824
; Patent No. 5932449
; GENERAL INFORMATION:
; APPLICANT: EMANUEL, PETER A.
; APPLICANT: BURANS, JAMES P.
; APPLICANT: VALDES, JAMES J.
; APPLICANT: MOHYEE, ELDEFRAWI E.
; TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: U.S. Army Chemical and Biological Defense
; ADDRESSEE: Command
; STREET: Office of the Chief Counsel, Bldg E4435
; CITY: Aberdeen Proving Ground
; STATE: MD
; COUNTRY: U.S.
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/792,824
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Biffoni, U. J.
; REGISTRATION NUMBER: 39,908
; REFERENCE/DOCKET NUMBER: DAM 431-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 410-671-1158
; TELEFAX: 410-671-2534
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-792-824-10
;
; Query Match 75.0%; Score 69; DB 2; Length 254;
; Best Local Similarity 75.0%; Pred. No. 0.0015;
; Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
;
QY 1 KIDPANGTKYDPIFQ 16
Db :|||||:|||||
72 RIDPANGNTEYDPKFQ 87

RESULT 17
US-08-792-824-10
; Sequence 10, Application US/08792824
; Patent No. 5932449
; GENERAL INFORMATION:
; APPLICANT: EMANUEL, PETER A.
; APPLICANT: BURANS, JAMES P.
; APPLICANT: VALDES, JAMES J.
; APPLICANT: MOHYEE, ELDEFRAWI E.
; TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: U.S. Army Chemical and Biological Defense
; ADDRESSEE: Command
; STREET: Office of the Chief Counsel, Bldg E4435
; CITY: Aberdeen Proving Ground
; STATE: MD
; COUNTRY: U.S.
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/792,824
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Biffoni, U. J.
; REGISTRATION NUMBER: 39,908
; REFERENCE/DOCKET NUMBER: DAM 431-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 410-671-1158
; TELEFAX: 410-671-2534
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-792-824-10
;
; Query Match 75.0%; Score 69; DB 2; Length 254;
; Best Local Similarity 75.0%; Pred. No. 0.0015;
; Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
;
QY 1 KIDPANGTKYDPIFQ 16
Db :|||||:|||||
72 RIDPANGNTEYDPKFQ 87

RESULT 18
US-08-792-824-13
; Sequence 13, Application US/08792824
; Patent No. 5932449
; GENERAL INFORMATION:
; APPLICANT: EMANUEL, PETER A.
; APPLICANT: BURANS, JAMES P.
; APPLICANT: VALDES, JAMES J.
```

; APPLICANT: MOHVEE, ELDEFRAWI E.  
; TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: U.S. Army Chemical and Biological Defense  
; ADDRESSEE: Command  
; STREET: Office of the Chief Counsel, Bldg E4435  
; CITY: Aberdeen Proving Ground  
; STATE: MD  
; COUNTRY: U.S.  
; ZIP: 21010-5423  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/792,824  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Biffoni, U. J.  
; REGISTRATION NUMBER: 39,908  
; REFERENCE/DOCKET NUMBER: DAM 431-96  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 410-671-1158  
; TELEFAX: 410-671-2534  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 254 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-792-824-13

Query Match 75.0%; Score 69; DB 2; Length 254;  
Best Local Similarity 75.0%; Pred. No. 0.0015;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16  
Db 72 RIDPANGTEYDPKQ 87

RESULT 19  
US-08-836-561-47  
; Sequence 47, Application US/08836561  
; Patent No. 6018032  
; GENERAL INFORMATION:  
; APPLICANT: KOIKE, Masamichi  
; APPLICANT: FURUYA, Akiko  
; APPLICANT: NAKAMURA, Kazuyasu  
; APPLICANT: IIDA, Akihiko  
; APPLICANT: ANAZAWA, Hideharu  
; APPLICANT: HANAI, No. 6018032uo  
; APPLICANT: TAKATSU, Kiyoshi  
; TITLE OF INVENTION: Antibody Against Human Interleukin-5  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/836,561

; FILING DATE: 09-MAY-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 232384/95  
; FILING DATE: 11-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lawrence, III, Stanton T  
; REGISTRATION NUMBER: 25,736  
; REFERENCE/DOCKET NUMBER: 7005-115-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-836-561-47

Query Match 73.9%; Score 68; DB 3; Length 17;  
Best Local Similarity 76.5%; Pred. No. 0.00011;  
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFOA 17  
Db 1 RIDPANGTKSDPKFOA 17

RESULT 20  
US-08-646-265A-116  
; Sequence 116, Application US/08646265A  
; Patent No. 6214973  
; GENERAL INFORMATION:  
; APPLICANT: OHTOMO, Toshihiko  
; APPLICANT: SATO, Koh  
; APPLICANT: TSUCHIYA, Masayuki  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; TITLE OF INVENTION: MEDULLOBLASTOMA CELLS  
; NUMBER OF SEQUENCES: 132  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/646,265A  
; FILING DATE: 09-SEP-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/JP94/01763  
; FILING DATE: 19-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 5-291078  
; FILING DATE: 19-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WEGNER, Harold C.  
; REGISTRATION NUMBER: 25,258  
; REFERENCE/DOCKET NUMBER: 53466/184  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 116:

SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-265A-116

Query Match 73.9%; Score 68; DB 3; Length 17;  
Best Local Similarity 75.0%; Pred. No. 0.00011;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:||||:|||||  
Db 1 RIDPADGNTKYDPKFK 16

## RESULT 21

US-09-434-122-47  
Sequence 47, Application US/09434122  
Patent No. 6538111  
GENERAL INFORMATION:  
APPLICANT: KOIKE, Masamichi  
FURUYA, Akiko  
NAKAMURA, Kazuyasu  
IIDA, Akihiro  
ANAZAWA, Hideharu  
HANAI, No. 6538111uo  
TAKATSU, Kiyoshi  
TITLE OF INVENTION: Antibody Against Human Interleukin-5 Receptor Alpha Chain  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Fast-SEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/434,122  
FILING DATE: 05-NO. 6538111-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/836,561  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: JP 232384/95  
FILING DATE: 11-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lawrence, III, Stanton T  
REGISTRATION NUMBER: 25,736  
REFERENCE/DOCKET NUMBER: 7005-115-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 47:

Query Match 73.9%; Score 68; DB 4; Length 17;  
Best Local Similarity 76.5%; Pred. No. 0.00011;  
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 17  
:|||||:|||||  
Db 1 RIDPANGNTKSDPKFQ 17

## RESULT 22

US-08-646-265A-132  
Sequence 132, Application US/08646265A  
Patent No. 6214973  
GENERAL INFORMATION:  
APPLICANT: OHTOMO, Toshihiko  
APPLICANT: SATO, Koh  
APPLICANT: TSUCHIYA, Masayuki  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
MEDULLOBLASTOMA CELLS  
NUMBER OF SEQUENCES: 132  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,265A  
FILING DATE: 09-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/JP94/01763  
FILING DATE: 19-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-291078  
FILING DATE: 19-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/184  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 132:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 117 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-265A-132

Query Match 73.9%; Score 68; DB 3; Length 117;  
Best Local Similarity 75.0%; Pred. No. 0.00095;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:||||:|||||  
Db 50 RIDPADGNTKYDPKFK 65

## RESULT 23

US-08-646-265A-29  
Sequence 29, Application US/08646265A  
Patent No. 6214973  
GENERAL INFORMATION:  
APPLICANT: OHTOMO, Toshihiko  
APPLICANT: SATO, Koh  
APPLICANT: TSUCHIYA, Masayuki  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
MEDULLOBLASTOMA CELLS

```
;
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,265A
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01763
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-291078
; FILING DATE: 19-NOV-1993
; NAME: WEGNER, Harold C.
; ATTORNEY/AGENT INFORMATION:
; FILING DATE: 19-NOV-1993
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 136 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-646-265A-29

Query Match 73.9%; Score 68; DB 3; Length 136;
Best Local Similarity 75.0%; Pred. No. 0.0011;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 69 RIDPADGNTKYDPKFQ 84

RESULT 24
US-08-646-265A-99
; Sequence 99, Application US/08646265A
; Patent No. 6214973
; GENERAL INFORMATION:
; APPLICANT: OHTOMO, Toshihiko
; APPLICANT: SATO, Koh
; APPLICANT: TSUCHIYA, Masayuki
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
; TITLE OF INVENTION: MEDULLOBLASTOMA CELLS
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,265A
```

```
;
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01763
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-291078
; FILING DATE: 19-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 136 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-646-265A-99

Query Match 73.9%; Score 68; DB 3; Length 136;
Best Local Similarity 75.0%; Pred. No. 0.0011;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 69 RIDPADGNTKYDPKFQ 84

RESULT 25
US-08-836-561-31
; Sequence 31, Application US/08836561
; Patent No. 6018032
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Masamichi
; APPLICANT: FURUYA, Akiko
; APPLICANT: NAKAMURA, Kazuyasu
; APPLICANT: IIDA, Akihiro
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: HANAI, No. 6018032uo
; APPLICANT: TAKATSU, Kiyoshi
; TITLE OF INVENTION: Antibody Against Human Interleukin-5
; TITLE OF INVENTION: Receptor Alpha Chain
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,561
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 232384/95
; FILING DATE: 11-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lawrence, III, Stanton T
; REGISTRATION NUMBER: 25,736
; REFERENCE/DOCKET NUMBER: 7005-115-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
```



TELEFAX: 212-869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 137 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-836-561-31

Query Match 73.9%; Score 68; DB 3; Length 137;  
Best Local Similarity 76.5%; Pred. No. 0.0011;  
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17  
Db 69 RIDPANGTKSDPKFOA 85  
:|||||:|||||

RESULT 26  
US-09-434-122-31  
; Sequence 31, Application US/09434122  
; Patent No. 6538111

GENERAL INFORMATION:  
APPLICANT: KOIKE, Masamichi  
FURUYA, Akiro  
NAKAMURA, Kazuyasu  
IIDA, Akihiro  
ANAZAWA, Hideharu  
HANAI, No. 6538111uo  
TAKATSU, Kiyoshi  
TITLE OF INVENTION: Antibody Against Human Interleukin-5  
Receptor Alpha Chain  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/434,122  
FILING DATE: 05-NO. 6538111-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/836,561  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: JP 232384/95  
FILING DATE: 11-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lawrence, III, Stanton T  
REGISTRATION NUMBER: 25,736  
REFERENCE/DOCKET NUMBER: 7005-115-999

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 137 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-09-434-122-31

Query Match 73.9%; Score 68; DB 4; Length 137;  
Best Local Similarity 76.5%; Pred. No. 0.0011;  
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17  
Db 69 RIDPANGTKSDPKFOA 85  
:|||||:|||||

RESULT 27  
US-08-646-265A-109  
; Sequence 109, Application US/08646265A  
; Patent No. 6214573  
GENERAL INFORMATION:  
APPLICANT: OHTOMO, Toshihiko  
APPLICANT: SATO, Koh  
APPLICANT: TSUCHIYA, Masayuki  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
TITLE OF INVENTION: MEDULLOBLASTOMA CELLS  
NUMBER OF SEQUENCES: 132  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,265A  
FILING DATE: 09-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/JP94/01763  
FILING DATE: 19-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-291078  
FILING DATE: 19-NOV-1993

ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/184  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 109:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 269 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-265A-109

Query Match 73.9%; Score 68; DB 3; Length 269;  
Best Local Similarity 75.0%; Pred. No. 0.0024;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFO 16  
Db 72 RIDPADGNTKYDPKFQ 87  
:|||||:|||||

RESULT 28  
US-08-950-660-2  
; Sequence 2, Application US/08950660  
; Patent No. 5932214  
GENERAL INFORMATION:  
APPLICANT: Lobb, Roy R.; Burkly, Linda C.

```
; TITLE OF INVENTION: Treatment For Inflammatory Bowel Disease
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/950,660
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/456,124
; FILING DATE:
; APPLICATION NUMBER: US 08/373,857
; FILING DATE: 18-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/284,603
; FILING DATE: 11-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00924
; FILING DATE: 02-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,139
; FILING DATE: 12-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis (PLM)
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-031USCP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-950-660-2

Query Match 72.8%; Score 67; DB 2; Length 120;
Best Local Similarity 75.0%; Pred. No. 0.0014;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
:||||:|||||
Db 49 RIDPASGDTKYDPKFQ 64

RESULT 29
US-08-454-899G-6
; Sequence 6, Application US/08454899G
; Patent No. 6502503
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/08/454,899G
; CURRENT FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 6
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: heavy chain variable region
US-08-454-899G-6

Query Match 72.8%; Score 67; DB 4; Length 120;
Best Local Similarity 75.0%; Pred. No. 0.0014;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
:||||:|||||
Db 49 RIDPASGDTKYDPKFQ 64

RESULT 30
PCT-US93-00030-2
; Sequence 2, Application PC/TUS9300030
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; TITLE OF INVENTION: Treatment for Asthma
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00030
; FILING DATE: 19930112
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 92,307-A; D002 CIP PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US93-00030-2

Query Match 72.8%; Score 67; DB 5; Length 120;
Best Local Similarity 75.0%; Pred. No. 0.0014;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
:||||:|||||
Db 49 RIDPASGDTKYDPKFQ 64

RESULT 31
PCT-US93-00924-2
; Sequence 2, Application PC/TUS9300924
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; TITLE OF INVENTION: Treatment for Inflammatory Bowel Disease
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
```

STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: IL  
COUNTRY: US  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00924  
FILING DATE: 19930202  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 92,308-A; D003 CIP PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-00924-2

Query Match 72.8%; Score 67; DB 5; Length 120;  
Best Local Similarity 75.0%; Pred. No. 0.0014;  
Matches 12; Conservative 2; Mismatches 2; Indels

```
Qy      1 KIDPANGKTKYDPFIQ 16
        :|:|:|:|:|:|:|
Db     49 RIDPASGDTKYDPKFO 64
```

```

RESULT 32
US-08-822-830B-2
; Sequence 2, Application US/08822830B
; Patent No. 5871734
; GENERAL INFORMATION:
; APPLICANT: Lobbs, Roy R.; Burkly, Linda C.
; TITLE OF INVENTION: Treatment for Asthma with VLA-4 Blocking
; TITLE OF INVENTION: Agents
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,830B
; FILING DATE: 03-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/456,193
; FILING DATE: 31-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/374,331
; FILING DATE: 18-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,631
; FILING DATE: 12-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00030

```

? FILING DATE: 12-JAN-1993  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: 07/821,768  
? FILING DATE: 13-JAN 1992  
? ATTORNEY/AGENT INFORMATION:  
? NAME: Myers, Louis (PLM)  
? REGISTRATION NUMBER: 35,965  
? REFERENCE/DOCKET NUMBER: BGP-021USCNS  
? TELECOMMUNICATION INFORMATION:  
? TELEPHONE: (617) 227-7400  
? TELEFAX: (617) 742-4214  
? INFORMATION FOR SEQ ID NO: 2:  
? SEQUENCE CHARACTERISTICS:  
? LENGTH: 121 amino acids  
? TYPE: amino acid  
? TOPOLOGY: linear  
? MOLECULE TYPE: protein  
? US-08-822-830B-2

```
Query Match      72.8%; Score 67; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 0.0014;
Matches 12: Conservative 2; Mismatches 2; Indels
```

**Qy**            1 KIDPANGKTKYDPIFQ 16  
               :|||:|||||  
**Db**            50 RIDPASGDTKYDPKFQ 65

RESULT 33  
US-08-822-830B-13  
; Sequence 13, Application US/08822830B  
; Patent No. 5871734  
; GENERAL INFORMATION:  
; APPLICANT: Lobb, Roy R.; Burkly, Linda C.  
; TITLE OF INVENTION: Treatment for Asthma with VLA-4 Blocking  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA

```

02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,830B
FILING DATE: 03-MAR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/456,193
FILING DATE: 31-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/374,331
FILING DATE: 18-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,631
FILING DATE: 12-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00030
FILING DATE: 12-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/821,768
FILING DATE: 13-JAN 1992
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis (PLM)
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-021USCN
TELEPHONE (617)227-7400

```

```
; TELFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 121 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-822-830B-13

Query Match          72.8%; Score 67; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 0.0014;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 50 RIDPASGDTKYDPKFQ 65

RESULT 34
US-09-157-452B-2
; Sequence 2, Application US/09157452B
; Patent No. 6482409
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Burkly, Linda C.
; TITLE OF INVENTION: TREATMENT FOR INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 10274-004003
; CURRENT APPLICATION NUMBER: US/09/157,452B
; CURRENT FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 08/950,660
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: US 08/373,857
; PRIOR FILING DATE: 1995-01-18
; PRIOR APPLICATION NUMBER: US 08/284,603
; PRIOR FILING DATE: 1994-08-11
; PRIOR APPLICATION NUMBER: PCT/US93/00924
; PRIOR FILING DATE: 1993-02-02
; PRIOR APPLICATION NUMBER: US 07/835,139
; PRIOR FILING DATE: 1992-02-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-157-452B-2

Query Match          72.8%; Score 67; DB 4; Length 121;
Best Local Similarity 75.0%; Pred. No. 0.0014;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 50 RIDPASGDTKYDPKFQ 65

RESULT 35
US-09-157-452B-16
; Sequence 16, Application US/09157452B
; Patent No. 6482409
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Burkly, Linda C.
; TITLE OF INVENTION: TREATMENT FOR INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 10274-004003
; CURRENT APPLICATION NUMBER: US/09/157,452B
; CURRENT FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 08/950,660
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: US 08/373,857
; PRIOR FILING DATE: 1995-01-18
; PRIOR APPLICATION NUMBER: US 08/284,603
; PRIOR FILING DATE: 1994-08-11
```

```
; PRIOR APPLICATION NUMBER: PCT/US93/00924
; PRIOR FILING DATE: 1993-02-02
; PRIOR APPLICATION NUMBER: US 07/835,139
; PRIOR FILING DATE: 1992-02-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-157-452B-16

Query Match          72.8%; Score 67; DB 4; Length 121;
Best Local Similarity 75.0%; Pred. No. 0.0014;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 50 RIDPASGDTKYDPKFQ 65

RESULT 36
US-08-454-899G-47
; Sequence 47, Application US/08454899G
; Patent No. 6602503
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/08/454,899G
; CURRENT FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: heavy chain variable region
US-08-454-899G-47

Query Match          72.8%; Score 67; DB 4; Length 124;
Best Local Similarity 75.0%; Pred. No. 0.0015;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 50 RIDPASGDTKYDPKFQ 65

RESULT 37
US-08-463-128-4
; Sequence 4, Application US/08463128
; Patent No. 5695755
; GENERAL INFORMATION:
; APPLICANT: Papayannopoulou, Thalia (USA only)
; APPLICANT: Board of Regents, U.
; APPLICANT: Washington (except USA)
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
```

```
;
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,128
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-11-1993
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 143 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-463-128-4

Query Match 72.8%; Score 67; DB 1; Length 143;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16
:||||:|||||
Db 69 RIDPASGDTKYDPKFQ 84

RESULT 38
US-08-463-128-8
; Sequence 8, Application US/08463128
; Patent No. 5695755
; GENERAL INFORMATION:
; APPLICANT: Papayannopolou, Thalia (USA only)
; APPLICANT: Board of Regents, U.
; APPLICANT: Washington (except USA)
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,128
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-11-1993
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
```

```
;
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 143 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-463-128-8

Query Match 72.8%; Score 67; DB 1; Length 143;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16
:||||:|||||
Db 69 RIDPASGDTKYDPKFQ 84

RESULT 39
US-08-463-298-4
; Sequence 4, Application US/08463298
; Patent No. 5824304
; GENERAL INFORMATION:
; APPLICANT: Papayannopolou, Thalia (USA only)
; APPLICANT: Board of Regents, U.
; APPLICANT: Washington (except USA)
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,298
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-NOV-1993
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 143 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-463-298-4

Query Match 72.8%; Score 67; DB 2; Length 143;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

Qy 1 KIDPANGTKYDPIFQ 16  
:||||:|||||  
Db 69 RIDPASGDTKYDPKFQ 84

## RESULT 40

US-08-463-298-8  
; Sequence 8, Application US/08463298  
; Patent No. 5824304  
; GENERAL INFORMATION:  
; APPLICANT: Papavannopoulou, Thalia (USA only)  
; APPLICANT: Board of Regents, U.  
; APPLICANT: Washington (except USA)  
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM  
; TITLE OF INVENTION: CELLS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: C/O FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,298  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/11060  
; FILING DATE: 11-NOV-1993  
; APPLICATION NUMBER: US 07/977,702  
; FILING DATE: 13-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr., James F.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B173CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 596-9000  
; TELEFAX: (212) 596-9090  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 143 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-463-298-8

Query Match 72.8%; Score 67; DB 2; Length 143;  
Best Local Similarity 75.0%; Pred. No. 0.0017;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:||||:|||||  
Db 69 RIDPASGDTKYDPKFQ 84

Search completed: June 3, 2005, 16:03:55  
Job time : 39.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 16:01:46 ; Search time 123.121 Seconds

(without alignments)  
47.730 Million cell updates/sec

Title: US-10-089-452-22

Perfect score: 92

Sequence: 1 KIDPANGKTKYDPIFQA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1465611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 1465611

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	79.3	17	US-10-344-779-9	Sequence 9, Appli
2	73	79.3	17	US-10-738-809-9	Sequence 9, Appli
3	73	79.3	17	US-10-687-035-67	Sequence 67, Appli
4	73	79.3	93	US-10-179-851-7	Sequence 7, Appli
5	73	79.3	125	US-10-763-424-70	Sequence 70, Appli
6	73	79.3	133	US-10-344-779-4	Sequence 4, Appli
7	73	79.3	133	US-10-738-809-4	Sequence 4, Appli
8	73	79.3	141	US-10-687-035-55	Sequence 55, Appli
9	72	78.3	110	US-10-179-851-2	Sequence 2, Appli
10	72	78.3	110	US-10-179-851-3	Sequence 3, Appli
11	71	77.2	123	US-10-763-424-59	Sequence 59, Appli
12	71	77.2	123	US-10-763-424-61	Sequence 61, Appli
13	71	77.2	142	US-10-763-424-67	Sequence 67, Appli

14	71	77.2	144	17	US-10-763-424-54	Sequence 54, Appli
15	70	76.1	109	15	US-10-460-471-28	Sequence 28, Appli
16	69	75.0	17	17	US-10-855-013-30	Sequence 30, Appli
17	69	75.0	117	17	US-10-855-013-27	Sequence 27, Appli
18	68	73.9	17	10	US-09-749-873-116	Sequence 116, App
19	68	73.9	17	14	US-10-283-349-47	Sequence 47, Appli
20	68	73.9	117	10	US-09-749-873-132	Sequence 132, App
21	68	73.9	136	10	US-09-749-873-99	Sequence 99, Appli
22	68	73.9	136	10	US-09-749-873-31	Sequence 31, Appli
23	68	73.9	137	14	US-10-283-349-31	Sequence 109, App
24	68	72.8	269	10	US-09-749-873-109	Sequence 1, Appli
25	67	72.8	120	14	US-10-095-496-1	Sequence 6, Appli
26	67	72.8	120	14	US-10-428-662-6	Sequence 2, Appli
27	67	72.8	120	17	US-10-875-282-2	Sequence 2, Appli
28	67	72.8	120	17	US-10-763-424-2	Sequence 2, Appli
29	67	72.8	120	17	US-10-875-469-2	Sequence 2, Appli
30	67	72.8	121	14	US-10-252-978-2	Sequence 16, Appli
31	67	72.8	121	14	US-10-252-978-16	Sequence 47, Appli
32	67	72.8	124	14	US-10-428-662-47	Sequence 14, Appli
33	67	72.8	124	17	US-10-875-282-14	Sequence 14, Appli
34	67	72.8	124	17	US-10-763-424-14	Sequence 14, Appli
35	67	72.8	124	17	US-10-875-469-14	Sequence 14, Appli
36	67	72.8	143	14	US-10-428-662-33	Sequence 29, Appli
37	67	72.8	143	14	US-10-428-662-33	Sequence 33, Appli
38	67	72.8	143	14	US-10-428-662-39	Sequence 39, Appli
39	67	72.8	143	14	US-10-428-662-43	Sequence 43, Appli
40	67	72.8	143	14	US-10-428-662-51	Sequence 51, Appli
41	67	72.8	143	14	US-10-428-662-55	Sequence 55, Appli
42	67	72.8	143	14	US-10-428-662-69	Sequence 69, Appli
43	67	72.8	143	17	US-10-875-282-6	Sequence 6, Appli
44	67	72.8	143	17	US-10-875-282-10	Sequence 10, Appli
45	67	72.8	143	17	US-10-875-282-12	Sequence 12, Appli

#### ALIGNMENTS

#### RESULT 1

US-10-344-779-9  
; Sequence 9, Application US/10344779  
; Publication No. US2003021106A1  
; GENERAL INFORMATION:  
; APPLICANT: SWEET, RAYMOND W.  
; APPLICANT: TORNETTA, MARK A.  
; APPLICANT: TRUNER, ALEMESEGED  
; APPLICANT: WAITAM, TREVOR A.  
; TITLE OF INVENTION: ANTI-RANK LIGAND MONOCLONAL ANTIBODIES  
; FILE REFERENCE: GP50033  
; CURRENT APPLICATION NUMBER: US/10/344,779  
; CURRENT FILING DATE: 2003-02-17  
; PRIOR APPLICATION NUMBER: PCT/US01/26161  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,524  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 60/230,639  
; PRIOR FILING DATE: 2000-09-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-344-779-9

Query Match 79.3%; Score 73; DB 15; Length 17;

Best Local Similarity 81.2%; Pred. No. 2.9e-05; Indels 0; Gaps 0;

Matches 13; Conservative 1; Mismatches 2;

Oy 1 KIDPANGKTKYDPIFQ 16

Db 1 RIDPANGTKYDPKFK 16

```
RESULT 2
US-10-738-809-9
; Sequence 9, Application US/10738809
; Publication No. US20040171117A1
; GENERAL INFORMATION:
; APPLICANT: SWEET, RAYMOND W.
; APPLICANT: TORNETTA, MARK A.
; APPLICANT: TRUNER, ALEMESEGED
; APPLICANT: WATTAM, TREVOR A.
; TITLE OF INVENTION: ANTI-RANK LIGAND MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: USEFUL IN TREATMENT OF RANK LIGAND MEDIATED DISORDERS
; FILE REFERENCE: GP50033
; CURRENT APPLICATION NUMBER: US/10/738,809
; CURRENT FILING DATE: 2003-12-17
; PRIOR APPLICATION NUMBER: US/10/344,779
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: PCT/US01/26161
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,524
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 60/230,639
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-738-809-9

Query Match          79.3%; Score 73; DB 16; Length 17;
Best Local Similarity 81.2%; Pred. No. 2.9e-05;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16
Db 1 RIDPANGNTKYDPKFQ 16

RESULT 3
US-10-687-035-67
; Sequence 67, Application US/10687035
; Publication No. US20050064518A1
; GENERAL INFORMATION:
; APPLICANT: Albone, Earl F.
; APPLICANT: Soltis, Daniel A.
; TITLE OF INVENTION: ANTIBODIES THAT BIND CELL-ASSOCIATED
; TITLE OF INVENTION: CA 125/6772P AND METHODS OF USE THEREOF
; FILE REFERENCE: 6750-214-999
; CURRENT APPLICATION NUMBER: US/10/687,035
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: 60/485,986
; PRIOR FILING DATE: 2003-07-10
; PRIOR APPLICATION NUMBER: 60/418,828
; PRIOR FILING DATE: 2003-10-12
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 16H9 VH2 CDR
US-10-687-035-67

Query Match          79.3%; Score 73; DB 17; Length 17;
Best Local Similarity 81.2%; Pred. No. 2.9e-05;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16
Db 1 RIDPANGNTKYDPKFQ 16
```

```
RESULT 4
US-10-179-851-7
; Sequence 7, Application US/10179851
; Publication No. US20030017518A1
; GENERAL INFORMATION:
; APPLICANT: Lam et al.
; APPLICANT: Lam, Joseph S
; APPLICANT: Zhao, Xin S
; TITLE OF INVENTION: NON-RADIOACTIVE ASSAY OF LIPOPOLYSACCHARIDE KINASES
; FILE REFERENCE: BEP 3004
; CURRENT APPLICATION NUMBER: US/10/179,851
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: US 60/300,420
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-179-851-7

Query Match          79.3%; Score 73; DB 14; Length 93;
Best Local Similarity 81.2%; Pred. No. 0.0002;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16
Db 45 RIDPANGNTKYDPKFQ 60

RESULT 5
US-10-763-424-70
; Sequence 70, Application US/10763424
; Publication No. US20050069541A1
; GENERAL INFORMATION:
; APPLICANT: Karlik, Stephen J.
; APPLICANT: Pleiss, Michael A.
; APPLICANT: Konradi, Andrei W.
; APPLICANT: Grant, Francine S.
; APPLICANT: Semko, Christopher M.
; APPLICANT: Dessen, Darren B.
; APPLICANT: Messersmith, Elizabeth
; APPLICANT: Freedman, Stephen
; APPLICANT: Vednock, Ted
; TITLE OF INVENTION: Composition for and Treatment of Demyelinating Diseases
; TITLE OF INVENTION: and Paralysis By Administration of Remyelinating Agents
; FILE REFERENCE: 034008-061
; CURRENT APPLICATION NUMBER: US/10/763,424
; CURRENT FILING DATE: 2004-01-26
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 106, 120
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-763-424-70

Query Match          79.3%; Score 73; DB 17; Length 125;
Best Local Similarity 81.2%; Pred. No. 0.00027;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16
Db 50 RIDPANGNTKYDPKFQ 65
```



RESULT 6  
US-10-344-779-4  
; Sequence 4, Application US/10344779  
; Publication No. US20030211106A1  
; GENERAL INFORMATION:  
; APPLICANT: SWEET, RAYMOND W.  
; APPLICANT: TORNETTA, MARK A.  
; APPLICANT: TRUNEH, ALEMSEGED  
; APPLICANT: WATAM, TREVOR A.  
; TITLE OF INVENTION: ANTI-RANK LIGAND MONOCLONAL ANTIBODIES  
; FILE REFERENCE: GP50033  
; CURRENT FILING DATE: 2003-02-17  
; PRIOR APPLICATION NUMBER: PCT/US01/26161  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,524  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 60/230,639  
; PRIOR FILING DATE: 2000-09-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 133  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-344-779-4

Query Match 79.3%; Score 73; DB 15; Length 133;  
Best Local Similarity 81.2%; Pred. No. 0.00029;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16  
:||||| ||||| ||  
Db 50 RIDPANGTKYDPKFQ 65

RESULT 7  
US-10-738-809-4  
; Sequence 4, Application US/10738809  
; Publication No. US20040171117A1  
; GENERAL INFORMATION:  
; APPLICANT: SWEET, RAYMOND W.  
; APPLICANT: TORNETTA, MARK A.  
; APPLICANT: TRUNEH, ALEMSEGED  
; APPLICANT: WATAM, TREVOR A.  
; TITLE OF INVENTION: ANTI-RANK LIGAND MONOCLONAL ANTIBODIES  
; FILE REFERENCE: GP50033  
; CURRENT FILING DATE: 2003-12-17  
; PRIOR APPLICATION NUMBER: US/10/344,779  
; PRIOR FILING DATE: 2003-02-19  
; PRIOR APPLICATION NUMBER: PCT/US01/26161  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,524  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 60/230,639  
; PRIOR FILING DATE: 2000-09-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 133  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-738-809-4

Query Match 79.3%; Score 73; DB 16; Length 133;  
Best Local Similarity 81.2%; Pred. No. 0.00029;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16  
:||||| ||||| ||  
Db 50 RIDPANGTKYDPKFQ 65

RESULT 8  
US-10-687-035-55  
; Sequence 55, Application US/10687035  
; Publication No. US20050064518A1  
; GENERAL INFORMATION:  
; APPLICANT: Albone, Earl F.  
; APPLICANT: Soltis, Daniel A.  
; TITLE OF INVENTION: ANTIBODIES THAT BIND CELL-ASSOCIATED  
; FILE REFERENCE: 6750-214-999  
; CURRENT APPLICATION NUMBER: US/10/687,035  
; CURRENT FILING DATE: 2003-10-15  
; PRIOR APPLICATION NUMBER: 60/485,986  
; PRIOR FILING DATE: 2003-07-10  
; PRIOR APPLICATION NUMBER: 60/418,828  
; PRIOR FILING DATE: 2003-10-12  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 55  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: 16H9 heavy chain polypeptide variable region (16H9H)  
US-10-687-035-55

Query Match 79.3%; Score 73; DB 17; Length 141;  
Best Local Similarity 81.2%; Pred. No. 0.00031;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16  
:||||| ||||| ||  
Db 69 RIDPANGTKYDPKFQ 84

RESULT 9  
US-10-179-851-2  
; Sequence 2, Application US/10179851  
; Publication No. US20030017518A1  
; GENERAL INFORMATION:  
; APPLICANT: Lam et al.  
; APPLICANT: Zhao, Xin S  
; APPLICANT: Zhao, Xin S  
; TITLE OF INVENTION: NON-RADIOACTIVE ASSAY OF LIPOPOLYSACCHARIDE KINASES  
; FILE REFERENCE: BEP 3004  
; CURRENT APPLICATION NUMBER: US/10/179,851  
; CURRENT FILING DATE: 2002-08-20  
; PRIOR APPLICATION NUMBER: US 60/300,420  
; PRIOR FILING DATE: 2001-06-26  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 110  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-179-851-2

Query Match 78.3%; Score 72; DB 14; Length 110;  
Best Local Similarity 81.2%; Pred. No. 0.00035;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16  
:||||| ||||| ||  
Db 45 RIDPANGTKYDPKFQ 60

RESULT 10  
US-10-179-851-3  
; Sequence 3, Application US/10179851  
; Publication No. US20030017518A1

```
; GENERAL INFORMATION:
; APPLICANT: Lam et al.
; APPLICANT: Lam, Joseph S
; APPLICANT: Zhao, Xin S
; TITLE OF INVENTION: NON-RADIOACTIVE ASSAY OF LIPOPOLYSACCHARIDE KINASES
; FILE REFERENCE: BEP 3004
; CURRENT APPLICATION NUMBER: US/10/179,851
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: US 60/300,420
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(110)
; OTHER INFORMATION:
; US-10-179-851-3

Query Match      78.3%; Score 72; DB 14; Length 110;
Best Local Similarity 81.2%; Pred. No. 0.00035;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 KIDPANGKTKYDPIFQ 16
Db      45 RIDPANGHGTKYDPKFKQ 60

RESULT 11
US-10-763-424-59
; Sequence 59, Application US/10763424
; Publication No. US20050069541A1
; GENERAL INFORMATION:
; APPLICANT: Karlík, Stephen J.
; APPLICANT: Pleiss, Michael A.
; APPLICANT: Konradi, Andrei W.
; APPLICANT: Grant, Francine S.
; APPLICANT: Semko, Christopher M.
; APPLICANT: Dreszen, Darren B.
; APPLICANT: Messersmith, Elizabeth
; APPLICANT: Freedman, Stephen
; APPLICANT: Yednock, Ted
; TITLE OF INVENTION: Composition for and Treatment of Demyelinating Diseases
; FILE REFERENCE: 034008-061
; CURRENT APPLICATION NUMBER: US/10/763,424
; CURRENT FILING DATE: 2004-01-26
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-763-424-59

Query Match      77.2%; Score 71; DB 17; Length 123;
Best Local Similarity 81.2%; Pred. No. 0.00058;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 KIDPANGKTKYDPIFQ 16
Db      50 RIDPANGYTKYDPKFKQ 65

RESULT 12
US-10-763-424-61
; Sequence 61, Application US/10763424
; Publication No. US20050069541A1
; GENERAL INFORMATION:
; APPLICANT: Karlík, Stephen J.
; APPLICANT: Pleiss, Michael A.
```

```
; APPLICANT: Pleiss, Michael A.
; APPLICANT: Konradi, Andrei W.
; APPLICANT: Grant, Francine S.
; APPLICANT: Semko, Christopher M.
; APPLICANT: Dreszen, Darren B.
; APPLICANT: Messersmith, Elizabeth
; APPLICANT: Freedman, Stephen
; APPLICANT: Yednock, Ted
; TITLE OF INVENTION: Composition for and Treatment of Demyelinating Diseases
; FILE REFERENCE: 034008-061
; CURRENT APPLICATION NUMBER: US/10/763,424
; CURRENT FILING DATE: 2004-01-26
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-763-424-61

Query Match      77.2%; Score 71; DB 17; Length 123;
Best Local Similarity 81.2%; Pred. No. 0.00058;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 KIDPANGKTKYDPIFQ 16
Db      50 RIDPANGYTKYDPKFKQ 65

RESULT 13
US-10-763-424-67
; Sequence 67, Application US/10763424
; Publication No. US20050069541A1
; GENERAL INFORMATION:
; APPLICANT: Karlík, Stephen J.
; APPLICANT: Pleiss, Michael A.
; APPLICANT: Konradi, Andrei W.
; APPLICANT: Grant, Francine S.
; APPLICANT: Semko, Christopher M.
; APPLICANT: Dreszen, Darren B.
; APPLICANT: Messersmith, Elizabeth
; APPLICANT: Freedman, Stephen
; APPLICANT: Yednock, Ted
; TITLE OF INVENTION: Composition for and Treatment of Demyelinating Diseases
; FILE REFERENCE: 034008-061
; CURRENT APPLICATION NUMBER: US/10/763,424
; CURRENT FILING DATE: 2004-01-26
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-763-424-67

Query Match      77.2%; Score 71; DB 17; Length 142;
Best Local Similarity 81.2%; Pred. No. 0.00069;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 KIDPANGKTKYDPIFQ 16
Db      69 RIDPANGYTKYDPKFKQ 84

RESULT 14
US-10-763-424-54
; Sequence 54, Application US/10763424
; Publication No. US20050069541A1
; GENERAL INFORMATION:
; APPLICANT: Karlík, Stephen J.
; APPLICANT: Pleiss, Michael A.
```

; APPLICANT: Konradi, Andrei W.  
; APPLICANT: Grant, Francine S.  
; APPLICANT: Semko, Christopher M.  
; APPLICANT: Dessen, Darren B.  
; APPLICANT: Messersmith, Elizabeth  
; APPLICANT: Freedman, Stephen  
; APPLICANT: Yednock, Ted  
; TITLE OF INVENTION: Composition for and Treatment of Demyelinating Diseases  
; TITLE OF INVENTION: and Paralysis By Administration of Remyelating Agents  
; FILE REFERENCE: 034008-061  
; CURRENT APPLICATION NUMBER: US/10/763,424  
; CURRENT FILING DATE: 2004-01-26  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 54  
; LENGTH: 144  
; TYPE: PRT  
; ORGANISM: Mouse  
US-10-763-424-54

Query Match 77.2%; Score 71; DB 17; Length 144;  
Best Local Similarity 81.2%; Pred. No. 0.0007;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:|||||:|||||  
Db 69 RIDPANGTKYDPKPFQ 84

RESULT 15  
US-10-460-471-28  
; Sequence 28, Application US/10460471  
; Publication No. US20030206900A1  
; GENERAL INFORMATION:  
; APPLICANT: TERNYNCK, THERESA  
; APPLICANT: AVAMEAS, ALEXANDRE  
; APPLICANT: BUTTIN, GERARD  
; APPLICANT: AVAMEAS, STRAITIS  
; APPLICANT: SARON, MARIE-FRANCOISE  
; APPLICANT: BLONDEL, BRUNO  
; APPLICANT: COUDERC, THERESA  
; APPLICANT: MICHELSON, SUSAN  
; APPLICANT: ZIPETO, DONATO  
; TITLE OF INVENTION: VECTOR DERIVED FROM ANTIBODIES FOR TRANSFERRING SUBSTANCES INTO C  
; FILE REFERENCE: 0660-0166-0XCONT  
; CURRENT APPLICATION NUMBER: US/10/460,471  
; CURRENT FILING DATE: 2003-06-13  
; PRIOR APPLICATION NUMBER: US/09/497,997C  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: PCT/FR98/01740  
; PRIOR FILING DATE: 1998-08-04  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28  
; LENGTH: 109  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE  
US-10-460-471-28

Query Match 76.1%; Score 70; DB 15; Length 109;  
Best Local Similarity 81.2%; Pred. No. 0.00075;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:|||||:|||||  
Db 41 RIDPANGTKYAPKPFQ 56

RESULT 16  
US-10-855-013-30  
; Sequence 30, Application US/10855013

; Publication No. US20050025768A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc. et al.  
; TITLE OF INVENTION: ANTIBODIES AGAINST MONOCYTE CHEMOTACTIC PROTEINS  
; FILE REFERENCE: BGN-006PC  
; CURRENT APPLICATION NUMBER: US/10/855,013  
; CURRENT FILING DATE: 2004-05-27  
; PRIOR APPLICATION NUMBER: US 60/343391  
; PRIOR FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: US 60/383277  
; PRIOR FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: US 60/400469  
; PRIOR FILING DATE: 2002-08-01  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-855-013-30

Query Match 75.0%; Score 69; DB 17; Length 17;  
Best Local Similarity 75.0%; Pred. No. 0.00014;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:|||||:|||||  
Db 1 RIDPANGNTKFDPRFQ 16

RESULT 17  
US-10-855-013-27  
; Sequence 27, Application US/10855013  
; Publication No. US20050025768A1  
; GENERAL INFORMATION:  
; APPLICANT: Biogen, Inc. et al.  
; TITLE OF INVENTION: ANTIBODIES AGAINST MONOCYTE CHEMOTACTIC PROTEINS  
; FILE REFERENCE: BGN-006PC  
; CURRENT APPLICATION NUMBER: US/10/855,013  
; CURRENT FILING DATE: 2004-05-27  
; PRIOR APPLICATION NUMBER: US 60/343391  
; PRIOR FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: US 60/383277  
; PRIOR FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: US 60/400469  
; PRIOR FILING DATE: 2002-08-01  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-855-013-27

Query Match 75.0%; Score 69; DB 17; Length 117;  
Best Local Similarity 75.0%; Pred. No. 0.0012;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:|||||:|||||  
Db 50 RIDPANGNTKFDPRFQ 65

RESULT 18  
US-09-749-873-116  
; Sequence 116, Application US/09749873  
; Publication No. US20030023045A1  
; GENERAL INFORMATION:  
; APPLICANT: OHTOMO, Toshihiko  
; SATO, Koh  
; TSUCHIYA, Masayuki  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; MEDULLOBLASTOMA CELLS

```
;
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/749,873
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/646,265
; FILING DATE: 1996-09-09
; APPLICATION NUMBER: JP 5-291078
; FILING DATE: 19-NOV-1993
;
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
;
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 116:
;
; US-09-749-873-116
;
; Query Match 73.9%; Score 68; DB 10; Length 17;
; Best Local Similarity 75.0%; Pred. No. 0.0002;
; Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
;
; Qy 1 KIDPANGTKYDPIFQ 16
; Db 1 RIDPADGNTKYDPKFO 16
;
; RESULT 19
; US-10-283-349-47
; Sequence 47, Application US/10283349
; Publication No. US20030096977A1
; GENERAL INFORMATION:
; APPLICANT: KOIKE, Masamichi
; FURUYA, Akiko
; NAKAMURA, Kazuyasu
; IIDA, Akihiro
; ANAZAWA, Hideharu
; HANAI, No. US20030096977A1uo
; TAKATSU, Kiyoshi
; TITLE OF INVENTION: Antibody Against Human Interleukin-5
; Receptor Alpha Chain
;
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
;
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/283,349
; FILING DATE: 29-Oct-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,561
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: JP 232384/95
; FILING DATE: 11-SEP-1995
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Lawrence, III, Stanton T
; REGISTRATION NUMBER: 25,736
; REFERENCE/DOCKET NUMBER: 7005-115-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
;
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
;
; US-10-283-349-47
;
; Query Match 73.9%; Score 68; DB 14; Length 17;
; Best Local Similarity 76.5%; Pred. No. 0.0002;
; Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
;
; Qy 1 KIDPANGTKYDPIFOA 17
; Db 1 RIDPANGNTKSDPKFOA 17
;
; RESULT 20
; US-09-749-873-132
; Sequence 132, Application US/09749873
; Publication No. US20030023045A1
; GENERAL INFORMATION:
; APPLICANT: OHTOMO, Toshihiko
; SATO, Koh
; TSUCHIYA, Masayuki
;
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
; MEDULLOBLASTOMA CELLS
;
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/749,873
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/646,265
; FILING DATE: 1996-09-09
; APPLICATION NUMBER: JP 5-291078
; FILING DATE: 19-NOV-1993
;
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
```

REFERENCE/DOCKET NUMBER: 53466/184  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 132:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 117 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 132:  
US-09-749-873-132

Query Match 73.9%; Score 68; DB 10; Length 117;  
Best Local Similarity 75.0%; Pred. No. 0.0018;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:||||:|||||  
Db 50 RIDPADGNTKYPKFK 65

## RESULT 21

US-09-749-873-29  
; Sequence 29, Application US/09749873  
; Publication No. US20030023045A1  
; GENERAL INFORMATION:  
; APPLICANT: OHTOMO, Toshihiko  
; SATO, Koh  
; TSUCHIYA, Masayuki  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; MEDULLOBLASTOMA CELLS  
; NUMBER OF SEQUENCES: 132  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/749,873  
; FILING DATE: 29-Dec-2000  
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/646,265  
FILING DATE: 1996-09-09  
APPLICATION NUMBER: JP 5-291078  
FILING DATE: 19-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/184  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 136 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 29:  
US-09-749-873-29

Query Match 73.9%; Score 68; DB 10; Length 136;

Best Local Similarity 75.0%; Pred. No. 0.0021;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 KIDPANGTKYDPIFQ 16  
:||||:|||||  
Db 69 RIDPADGNTKYPKFK 84

## RESULT 22

US-09-749-873-99  
; Sequence 99, Application US/09749873  
; Publication No. US20030023045A1  
; GENERAL INFORMATION:  
; APPLICANT: OHTOMO, Toshihiko  
; SATO, Koh  
; TSUCHIYA, Masayuki  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; MEDULLOBLASTOMA CELLS  
; NUMBER OF SEQUENCES: 132  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/749,873  
; FILING DATE: 29-Dec-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/646,265  
; FILING DATE: 1996-09-09  
; APPLICATION NUMBER: JP 5-291078  
; FILING DATE: 19-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WEGNER, Harold C.  
; REGISTRATION NUMBER: 25,258  
; REFERENCE/DOCKET NUMBER: 53466/184  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 99:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 136 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 99:  
US-09-749-873-99

Query Match 73.9%; Score 68; DB 10; Length 136;  
Best Local Similarity 75.0%; Pred. No. 0.0021;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:||||:|||||  
Db 69 RIDPADGNTKYPKFK 84

## RESULT 23

US-10-283-349-31  
; Sequence 31, Application US/10283349  
; Publication No. US20030096977A1  
; GENERAL INFORMATION:  
; APPLICANT: KOIKE, Masamichi  
; FURUYA, Akiko

NAKAMURA, Kazuyasu  
IIDA, Akihiro  
ANAZAWA, Hideharu  
HANAI, No. US20030096977A1uo  
TAKATSU, Kiyoshi  
TITLE OF INVENTION: Antibody Against Human Interleukin-5  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/283,349  
FILING DATE: 29-Oct-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/836,561  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: JP 232384/95  
FILING DATE: 11-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lawrence, III, Stanton T  
REGISTRATION NUMBER: 25,736  
REFERENCE/DOCKET NUMBER: 7005-115-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 137 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 31:  
US-10-283-349-31  
Query Match 73.9%; Score 68; DB 14; Length 137;  
Best Local Similarity 76.5%; Pred. No. 0.0021;  
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 KIDPANGTKYDPIFOA 17  
Db 69 RIDPANGTKSDPKFOA 85  
RESULT 24  
US-09-749-873-109  
Sequence 109, Application US/09749873  
Publication No. US20030023045A1  
GENERAL INFORMATION:  
APPLICANT: OHTOMO, Toshihiko  
SATO, Koh  
TSUCHIYA, Masayuki  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
MEDULLOBLASTOMA CELLS  
NUMBER OF SEQUENCES: 132  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA

ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/749,873  
FILING DATE: 29-Dec-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/646,265  
FILING DATE: 1996-09-09  
APPLICATION NUMBER: JP 5-291078  
FILING DATE: 19-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/184  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 109:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 269 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 109:  
US-09-749-873-109  
Query Match 73.9%; Score 68; DB 10; Length 269;  
Best Local Similarity 75.0%; Pred. No. 0.0045;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 KIDPANGTKYDPIFO 16  
Db 72 RIDPADGNTKYDPKFQ 87  
RESULT 25  
US-10-095-496-1  
Sequence 1, Application US/10095496  
Publication No. US20030007969A1  
GENERAL INFORMATION:  
APPLICANT: Lobb, Roy  
APPLICANT: Allen, Andrew  
APPLICANT: Pusey, Charles  
APPLICANT: Biogen, Inc.  
TITLE OF INVENTION: Therapies for Chronic Renal Failure  
TITLE OF INVENTION: Using One or More Integrin Antagonists  
FILE REFERENCE: A082 US  
CURRENT APPLICATION NUMBER: US/10/095,496  
CURRENT FILING DATE: 2002-03-12  
PRIOR APPLICATION NUMBER: PCT/US00/25140  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: 60/153,826  
PRIOR FILING DATE: 1999-09-14  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 120  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HPI/2 heavy chain variable region  
US-10-095-496-1  
Query Match 72.8%; Score 67; DB 14; Length 120;  
Best Local Similarity 75.0%; Pred. No. 0.0027;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:||||:|||||  
Db 49 RIDPASGDTKYDPKFQ 64

## RESULT 26

US-10-428-662-6  
; Sequence 6, Application US/10428662  
; Publication No. US20030185819A1  
; GENERAL INFORMATION:  
; APPLICANT: Lobb, Roy R.  
; APPLICANT: Carr, Frank J.  
; APPLICANT: Tempest, Philip R.  
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES  
; FILE REFERENCE: 10274-007001  
; CURRENT APPLICATION NUMBER: US/10/428,662  
; CURRENT FILING DATE: 2003-05-02  
; PRIOR APPLICATION NUMBER: US/08/454,899G  
; PRIOR FILING DATE: 1995-05-31  
; PRIOR APPLICATION NUMBER: US 08/004,798  
; PRIOR FILING DATE: 1993-01-12  
; PRIOR APPLICATION NUMBER: PCT/US94/00266  
; PRIOR FILING DATE: 1994-01-07  
; NUMBER OF SEQ ID NOS: 109  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 120  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: heavy chain variable region  
US-10-428-662-6

Query Match 72.8%; Score 67; DB 14; Length 120;  
Best Local Similarity 75.0%; Pred. No. 0.0027;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:||||:|||||  
Db 49 RIDPASGDTKYDPKFQ 64

## RESULT 27

US-10-875-282-2  
; Sequence 2, Application US/10875282  
; Publication No. US20050065192A1  
; GENERAL INFORMATION:  
; APPLICANT: Vednock, Theodore A.  
; APPLICANT: Freedman, Stephen B.  
; APPLICANT: Lieberburg, Ivan  
; APPLICANT: Shopp, George  
; APPLICANT: Pleiss, Michael A.  
; APPLICANT: Konradi, Andrei W.  
; APPLICANT: Shopp, George  
; APPLICANT: Messersmith, Elizabeth  
; TITLE OF INVENTION: Methods and Compositions for Treating Rheumatoid Arthritis  
; FILE REFERENCE: 002010-856  
; CURRENT APPLICATION NUMBER: US/10/875,282  
; CURRENT FILING DATE: 2004-06-25  
; PRIOR APPLICATION NUMBER: US 60/482,211  
; PRIOR FILING DATE: 2003-06-25  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 120  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-875-282-2

Query Match 72.8%; Score 67; DB 17; Length 120;  
Best Local Similarity 75.0%; Pred. No. 0.0027;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:||||:|||||  
Db 49 RIDPASGDTKYDPKFQ 64

## RESULT 28

US-10-763-424-2  
; Sequence 2, Application US/10763424  
; Publication No. US20050069541A1  
; GENERAL INFORMATION:  
; APPLICANT: Karlik, Stephen J.  
; APPLICANT: Pleiss, Michael A.  
; APPLICANT: Konradi, Andrei W.  
; APPLICANT: Grant, Francine S.  
; APPLICANT: Semko, Christopher M.  
; APPLICANT: Dessen, Darren B.  
; APPLICANT: Messersmith, Elizabeth  
; APPLICANT: Freedman, Stephen  
; APPLICANT: Vednock, Ted  
; TITLE OF INVENTION: Composition for and Treatment of Demyelinating Diseases  
; TITLE OF INVENTION: and Paralysis By Administration of Remyelinating Agents  
; FILE REFERENCE: 034008-061  
; CURRENT APPLICATION NUMBER: US/10/763,424  
; CURRENT FILING DATE: 2004-01-26  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 120  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-763-424-2

Query Match 72.8%; Score 67; DB 17; Length 120;  
Best Local Similarity 75.0%; Pred. No. 0.0027;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:||||:|||||  
Db 49 RIDPASGDTKYDPKFQ 64

## RESULT 29

US-10-875-469-2  
; Sequence 2, Application US/10875469  
; Publication No. US20050074451A1  
; GENERAL INFORMATION:  
; APPLICANT: Vednock, Theodore A.  
; APPLICANT: Freedman, Stephen B.  
; APPLICANT: Lieberburg, Ivan  
; APPLICANT: Shopp, George  
; APPLICANT: Pleiss, Michael A.  
; APPLICANT: Konradi, Andrei W.  
; APPLICANT: Shopp, George  
; APPLICANT: Messersmith, Elizabeth  
; TITLE OF INVENTION: Methods and Compositions for Treating Rheumatoid Arthritis  
; FILE REFERENCE: 034008-072  
; CURRENT APPLICATION NUMBER: US/10/875,469  
; CURRENT FILING DATE: 2004-06-25  
; PRIOR APPLICATION NUMBER: US 60/482,211  
; PRIOR FILING DATE: 2003-06-25  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 120  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-875-469-2

Query Match 72.8%; Score 67; DB 17; Length 120;  
Best Local Similarity 75.0%; Pred. No. 0.0027;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy 1 KIDPANGTKYDPIFQ 16
Db 49 RIDPASGDTKYDPKFQ 64

RESULT 30
US-10-252-978-2
; Sequence 2, Application US/10252978
; Publication No. US20030095969A1
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Burkly, Linda C.
; TITLE OF INVENTION: TREATMENT FOR INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 10274-004003
; CURRENT APPLICATION NUMBER: US/10/252,978
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US/09/157,452
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 08/950,660
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: US 08/373,857
; PRIOR FILING DATE: 1995-01-18
; PRIOR APPLICATION NUMBER: US 08/284,603
; PRIOR FILING DATE: 1994-08-11
; PRIOR APPLICATION NUMBER: PCT/US93/00924
; PRIOR FILING DATE: 1993-02-02
; PRIOR APPLICATION NUMBER: US 07/835,139
; PRIOR FILING DATE: 1992-02-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-252-978-2

Query Match 72.8%; Score 67; DB 14; Length 121;
Best Local Similarity 75.0%; Pred. No. 0.0027;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 50 RIDPASGDTKYDPKFQ 65

RESULT 31
US-10-252-978-16
; Sequence 16, Application US/10252978
; Publication No. US20030095969A1
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Burkly, Linda C.
; TITLE OF INVENTION: TREATMENT FOR INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 10274-004003
; CURRENT APPLICATION NUMBER: US/10/252,978
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US/09/157,452
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 08/950,660
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: US 08/373,857
; PRIOR FILING DATE: 1995-01-18
; PRIOR APPLICATION NUMBER: US 08/284,603
; PRIOR FILING DATE: 1994-08-11
; PRIOR APPLICATION NUMBER: PCT/US93/00924
; PRIOR FILING DATE: 1993-02-02
; PRIOR APPLICATION NUMBER: US 07/835,139
; PRIOR FILING DATE: 1992-02-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 121
; TYPE: PRT
US-10-252-978-16

Query Match 72.8%; Score 67; DB 14; Length 121;
Best Local Similarity 75.0%; Pred. No. 0.0027;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 50 RIDPASGDTKYDPKFQ 65

RESULT 32
US-10-428-662-47
; Sequence 47, Application US/10428662
; Publication No. US20030185819A1
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/10/428,662
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US/08/454,899G
; PRIOR FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: heavy chain variable region
US-10-428-662-47

Query Match 72.8%; Score 67; DB 14; Length 124;
Best Local Similarity 75.0%; Pred. No. 0.0028;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
Db 50 RIDPASGDTKYDPKFQ 65

RESULT 33
US-10-875-282-14
; Sequence 14, Application US/10875282
; Publication No. US20050065192A1
; GENERAL INFORMATION:
; APPLICANT: Yednock, Theodore A.
; APPLICANT: Freedman, Stephen B.
; APPLICANT: Lieberburg, Ivan
; APPLICANT: Shopp, George
; APPLICANT: Pleiss, Michael A.
; APPLICANT: Konradi, Andrei W.
; APPLICANT: Shopp, George
; APPLICANT: Messersmith, Elizabeth
; TITLE OF INVENTION: Methods and Compositions for Treating Rheumatoid Arthritis
; FILE REFERENCE: 002010-856
; CURRENT APPLICATION NUMBER: US/10/875,282
; CURRENT FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/482,211
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 124
; TYPE: PRT
```



```

; ORGANISM: Homo sapiens
US-10-875-282-14
Query Match      72.8%; Score 67; DB 17; Length 124;
Best Local Similarity 75.0%; Pred. No. 0.0028;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16
Db 50 RIDPASGDTKYDPKRF 65

RESULT 34
US-10-763-424-14
; Sequence 14, Application US/10763424
; Publication No. US20050069541A1
; GENERAL INFORMATION:
; APPLICANT: Karlík, Stephen J.
; APPLICANT: Pleiss, Michael A.
; APPLICANT: Konradi, Andrei W.
; APPLICANT: Grant, Francine S.
; APPLICANT: Semko, Christopher M.
; APPLICANT: Dressen, Darren B.
; APPLICANT: Messersmith, Elizabeth
; APPLICANT: Freedman, Stephen
; APPLICANT: Yednock, Ted
; TITLE OF INVENTION: Composition for and Treatment of Demyelinating Diseases
; TITLE OF INVENTION: and Paralysis By Administration of Remyelating Agents
; FILE REFERENCE: 034008-061
; CURRENT APPLICATION NUMBER: US/10/763,424
; CURRENT FILING DATE: 2004-01-26
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-763-424-14

Query Match      72.8%; Score 67; DB 17; Length 124;
Best Local Similarity 75.0%; Pred. No. 0.0028;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16
Db 50 RIDPASGDTKYDPKRF 65

RESULT 35
US-10-875-469-14
; Sequence 14, Application US/10875469
; Publication No. US20050074451A1
; GENERAL INFORMATION:
; APPLICANT: Yednock, Theodore A.
; APPLICANT: Freedman, Stephen B.
; APPLICANT: Lieberburg, Ivan
; APPLICANT: Shopp, George
; APPLICANT: Pleiss, Michael A.
; APPLICANT: Konradi, Andrei W.
; APPLICANT: Shopp, George
; APPLICANT: Messersmith, Elizabeth
; TITLE OF INVENTION: Methods and Compositions for Treating Rheumatoid Arthritis
; FILE REFERENCE: 034008-072
; CURRENT APPLICATION NUMBER: US/10/875,469
; CURRENT FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/482,211
; PRIOR FILING DATE: 2003-06-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; ORGANISM: Artificial Sequence
; FEATURE:

US-10-875-469-14
Query Match      72.8%; Score 67; DB 17; Length 124;
Best Local Similarity 75.0%; Pred. No. 0.0028;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16
Db 50 RIDPASGDTKYDPKRF 65

RESULT 36
US-10-428-662-29
; Sequence 29, Application US/10428662
; Publication No. US20030185819A1
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/10/428,662
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US/08/454,899G
; PRIOR FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: heavy chain variable region
US-10-428-662-29

Query Match      72.8%; Score 67; DB 14; Length 143;
Best Local Similarity 75.0%; Pred. No. 0.0032;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIQF 16
Db 69 RIDPASGDTKYDPKRF 84

RESULT 37
US-10-428-662-33
; Sequence 33, Application US/10428662
; Publication No. US20030185819A1
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/10/428,662
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US/08/454,899G
; PRIOR FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: heavy chain variable region
US-10-428-662-33

Query Match          72.8%; Score 67; DB 14; Length 143;
Best Local Similarity 75.0%; Pred. No. 0.0032;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
   :|||:|||||
Db 69 RIDPASGDTKYDPKFQ 84

RESULT 38
US-10-428-662-39
; Sequence 39, Application US/10428662
; Publication No. US20030185819A1
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/10/428,662
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US/08/454,899G
; PRIOR FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: heavy chain variable region
US-10-428-662-39

Query Match          72.8%; Score 67; DB 14; Length 143;
Best Local Similarity 75.0%; Pred. No. 0.0032;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
   :|||:|||||
Db 69 RIDPASGDTKYDPKFQ 84

RESULT 39
US-10-428-662-43
; Sequence 43, Application US/10428662
; Publication No. US20030185819A1
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/10/428,662
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US/08/454,899G
; PRIOR FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: heavy chain variable region
US-10-428-662-43

Query Match          72.8%; Score 67; DB 14; Length 143;
Best Local Similarity 75.0%; Pred. No. 0.0032;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
   :|||:|||||
Db 69 RIDPASGDTKYDPKFQ 84

RESULT 40
US-10-428-662-51
; Sequence 51, Application US/10428662
; Publication No. US20030185819A1
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/10/428,662
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US/08/454,899G
; PRIOR FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: heavy chain variable region
US-10-428-662-51

Query Match          72.8%; Score 67; DB 14; Length 143;
Best Local Similarity 75.0%; Pred. No. 0.0032;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16
   :|||:|||||
Db 69 RIDPASGDTKYDPKFQ 84

Search completed: June 3, 2005, 16:24:41
Job time : 124.121 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 3, 2005, 15:48:25 ; Search time 29.8788 Seconds  
(without alignments)  
54.744 Million cell updates/sec

Title: US-10-089-452-22

Perfect score: 92

Sequence: 1 KIDPANGTKYDPIFQA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.\*

2: Pir2.\*

3: Pir3.\*

4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	73	79.3	99	D37262	Ig heavy chain V r
2	73	79.3	120	S03471	Ig heavy chain V-D
3	73	79.3	122	S06823	Ig heavy chain V r
4	73	79.3	123	PH1403	Ig heavy chain V r
5	69	75.0	85	E37262	Ig heavy chain V r
6	68	73.9	99	C37262	Ig heavy chain V r
7	68	73.9	117	S17586	Ig heavy chain V r
8	68	73.9	221	S49220	Ig gamma-1 chain -
9	67	72.8	115	S03482	Ig heavy chain V-D
10	67	72.8	178	S29594	Ig gamma chain (WM
11	66	71.7	107	A27646	Ig heavy chain V r
12	65	70.7	107	PH1013	Ig heavy chain V r
13	65	70.7	108	PH1012	Ig heavy chain V r
14	63	68.5	114	A47271	nitrophenyl phosph
15	63	68.5	116	S24289	Ig gamma chain V r
16	62	67.4	268	A56446	Ig heavy chain V r
17	59	64.1	137	S52445	Ig heavy chain V r
18	56	60.9	135	PI0246	Ig heavy chain V r
19	55	59.8	120	S03484	Ig heavy chain V-D
20	54	58.7	902	S54495	probable carrier p
21	53	57.6	82	A36025	Ig heavy chain V r
22	49	53.3	98	PH1144	Ig heavy chain V r
23	48	52.2	138	S24246	Ig mu heavy chain
24	48	52.2	1312	1 BNEYDL	RAD50 protein - ye
25	47	51.1	98	PH1429	Ig heavy chain V r
26	47	51.1	101	S12431	Ig heavy chain V r
27	47	51.1	115	PH1560	Ig heavy chain V r
28	47	51.1	122	PH1426	Ig heavy chain V r
29	47	51.1	137	PH1562	Ig heavy chain V r

30	47	51.1	138	2	PH1564	Ig heavy chain V r
31	47	51.1	138	2	PH1565	Ig heavy chain V r
32	47	51.1	144	2	PH1583	Ig heavy chain V r
33	47	51.1	147	2	PH1561	Ig heavy chain V r
34	46	50.0	132	2	PH1427	Ig heavy chain V r
35	46	50.0	249	2	S41374	single chain Fv an
36	46	50.0	507	2	H69186	conserved hypother
37	46	50.0	673	2	AC0238	heamin storage aya
38	46	50.0	673	2	T47006	hypothetical prote
39	45	48.9	98	2	PH1139	Ig heavy chain V r
40	45	48.9	118	2	S25174	Ig heavy chain V r
41	45	48.9	131	2	A27472	Ig heavy chain pre
42	45	48.9	136	2	S04576	Ig heavy chain pre
43	45	48.9	141	2	S31685	Ig heavy chain V r
44	45	48.9	231	2	T05241	hypothetical prote
45	45	48.9	288	2	S29690	Ig heavy chain VDJ

## ALIGNMENTS

## RESULT 1

D37262

Ig heavy chain V region (2E5) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 30-Aug-1991 #sequence\_revision 30-Aug-1991 #text\_change 21-Jan-2000

C:Accession: D37262

R:Goshorn, S.C.; Retzel, E.; Jemerson, R.

J. Biol. Chem. 266, 2134-2142, 1991

A:Title: Common structural features among monoclonal antibodies binding the same antigen

A:Reference number: A38601; MUID:91115823; PMID:1703527

A:Accession: D37262

A&gt;Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-99 &lt;GOS&gt;

A:Cross-references: GB:M57990; NID:g195046; PIDN:AAA63328.1; PID:g195047

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

P:7-90/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 79.3%; Score 73; DB 2; Length 99;

Best Local Similarity 81.2%; Pred. No. 8.8e-05;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16

Db :||||| ||||| ||

42 RIDPANGTKYDKPKFQ 57

## RESULT 2

S03471

Ig heavy chain V-D-J region (hybridoma G5 Bb 2.2) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Jun-2000

C:Accession: S03471; S07453

R:Rocca-Serra, J.; Matthes, H.W.; Kaartinen, M.; Milstein, C.; Theze, J.; Fougereau, M.

EMBO J. 2, 867-872, 1983

A:Title: Analysis of antibody diversity: V-D-J mRNA nucleotide sequence of four anti-GAT

hypervariable regions.

A:Reference number: S03471; MUID:84057768; PMID:6416834

A:Accession: S03471

A:Molecule type: mRNA

A:Residues: 7-120 &lt;ROC1&gt;

A:Cross-references: EMBL:X01820; NID:g51833; PIDN:CAA25962.1; PID:g1333983

A&gt;Note: this sequence was determined from the differentiated gene

R:Rocca-Serra, J.; Mazie, J.C.; Moinier, D.; Leclercq, L.; Somme, G.; Theze, J.; Fougere

J. Immunol. 129, 2554-2558, 1982

A:Title: The limited diversity of the mouse gamma-chains anti-GAT repertoire does not se

A:Reference number: S07453; MUID:83058021; PMID:6815271

A:Accession: S07453

A:Molecule type: protein

A:Residues: 1-43 &lt;ROC2&gt;

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin  
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 79.3%; Score 73; DB 2; Length 120;  
Best Local Similarity 81.2%; Pred. No. 0.00011;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:|||||:|||||  
Db 50 RIDPANGTKYDPKFQ 65

## RESULT 3

Ig heavy chain V region (clone IIC) - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 21-Jan-2000  
C;Accession: S06823  
R;Miller III, A.; Glasel, J.A.  
J. Mol. Biol. 209, 763-778, 1989  
A;Title: Comparative sequence and immunochemical analyses of murine monoclonal anti-morp  
A;Reference number: S06815; MUID:90064531; PMID:2555519  
A;Accession: S06823  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-122 <MIL>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-99/Domain: immunoglobulin homology <IMM>

Query Match 79.3%; Score 73; DB 2; Length 122;  
Best Local Similarity 81.2%; Pred. No. 0.00011;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:|||||:|||||  
Db 51 RIDPANGTKYDPKFQ 66

## RESULT 4

Ig heavy chain V region (clone micro m- 46-6, 46-12) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Jan-2000  
C;Accession: PH1403  
R;Shirasawa, T.; Miyazoe, I.; Hagiwara, S.; Kimoto, H.; Shigemoto, K.; Taniguchi, M.; Ta  
J. Exp. Med. 176, 1209-1214, 1992  
A;Title: Heavy chain variable (VH) region diversity generated by VH gene replacement in  
ia virus.  
A;Reference number: PH1403; MUID:93018837; PMID:1402663  
A;Accession: PH1403  
A;Molecule type: DNA  
A;Residues: 1-123 <SHI>  
C;Genetics:  
A;Introns: 16/1  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;35-118/Domain: immunoglobulin homology <IMM>

Query Match 79.3%; Score 73; DB 2; Length 123;  
Best Local Similarity 81.2%; Pred. No. 0.00011;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16  
:|||||:|||||  
Db 70 RIDPANGTKYDPKFQ 85

## RESULT 5

E37262  
Ig heavy chain V region (6H2) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Aug-1991 #sequence\_revision 30-Aug-1991 #text\_change 09-May-1997

## C;Accession: E37262

R;Goshorn, S.C.; Retzel, E.; Jemmerson, R.

J. Biol. Chem. 266, 2134-2142, 1991

A;Title: Common structural features among monoclonal antibodies binding the same antigen

A;Reference number: A38601; MUID:91115823; PMID:1703527

A;Accession: E37262

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-85 <GOS>

A;Cross-references: GB:M57991

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

Query Match 75.0%; Score 69; DB 2; Length 85;

Best Local Similarity 75.0%; Pred. No. 0.00034;

Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16

:|||||:|||||

Db 28 RIDPANGTRYDPKFQ 43

## RESULT 6

C37262  
Ig heavy chain V region (1F5) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 30-Aug-1991 #sequence\_revision 30-Aug-1991 #text\_change 21-Jan-2000

C;Accession: C37262

R;Goshorn, S.C.; Retzel, E.; Jemmerson, R.

J. Biol. Chem. 266, 2134-2142, 1991

A;Title: Common structural features among monoclonal antibodies binding the same antigen

A;Reference number: A38601; MUID:91115823; PMID:1703527

A;Accession: C37262

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-99 <GOS>

A;Cross-references: GB:M57989; NID:G194534; PIDN:AAA63320.1; PID:G194535

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;7-90/Domain: immunoglobulin homology <IMM>

Query Match 73.9%; Score 68; DB 2; Length 99;

Best Local Similarity 75.0%; Pred. No. 0.00059;

Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFQ 16

:|||||:|||||

Db 42 RIDPANGTKYDPIFQ 57

## RESULT 7

S17586

Ig heavy chain V region (E8) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 21-Jan-2000

C;Accession: S17586

R;Mylvaganam, S.E.; Paterson, Y.; Kaiser, K.; Bowdish, K.; Getzoff, E.D.

J. Mol. Biol. 221, 455-462, 1991

A;Title: Biochemical implications from the variable gene sequences of an anti-cytochrome

forms.

A;Reference number: S17586; MUID:92015240; PMID:1656053

A;Accession: S17586

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-117 <MYL>

A;Cross-references: EMBL:X60683; NID:G51820; PIDN:CAA43095.1; PID:G51821

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 73.9%; Score 68; DB 2; Length 117;

Best Local Similarity 75.0%; Pred. No. 0.0007;

Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFQ 16  
:|||||:|||||  
Db 50 RIDPANGTKYDPRFQ 65

## RESULT 8

S49220  
Ig gamma-1 chain - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 19-Mar-1997 #sequence\_revision 01-Aug-1997 #text\_change 21-Jan-2000  
C/Accession: S49220  
R/Kipp, B.; Becker, W.P.; Schlaak, M.M.  
submitted to the EMBL Data Library, September 1994  
A/Description: Cloning and expression of a recombinant mouse Fab-fragment recognizing a  
A/Reference number: S49220  
A/Accession: S49220  
A/Molecule type: mRNA  
A/Residues: 1-221 <KIP>  
A/Cross-references: EMBL:Z37502; NID:g541778; PIDN:CAA85732.1; PID:g541779  
A/Experimental source: strain Balb/c  
C/Superfamily: immunoglobulin C region; immunoglobulin homology  
C/Keywords: immunoglobulin  
F:1-120/Domain: V region #status predicted <VRG>  
F:121-221/Domain: C region #status predicted <CRG>  
F:139-203/Domain: immunoglobulin homology <IMW>

Query Match 73.9%; Score 68; DB 2; Length 221;  
Best Local Similarity 75.0%; Pred. No. 0.0014;  
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFQ 16  
:|||||:|||||  
Db 50 RIDPANGTKYDPRFQ 65

## RESULT 9

S03482  
Ig heavy chain V-D-J region (hybridoma G8 Ad 3.8) - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
A/Variety: strain BALB/c  
C/Date: 26-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 21-Jan-2000  
C/Accession: S03482; S07453  
R/Rocca-Serra, J.; Matthes, H.W.; Kaartinen, M.; Mäntelä, C.; Theze, J.; Fougereau, M.  
EMBO J. 2, 867-872, 1983  
A/Title: Analysis of antibody diversity: V-D-J mRNA nucleotide sequence of four anti-GAT  
hypervariable regions.

A/Reference number: S03471; MUID:84057768; PMID:6416834

A/Accession: S03482

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 10-115 <ROCI>

A/Cross-references: EMBL:X03219

A/Note: This sequence was determined from the differentiated gene

R/Rocca-Serra, J.; Mazie, J.C.; Moinier, D.; Leclercq, L.; Somme, G.; Theze, J.; Fougere

J. Immunol. 129, 2554-2558, 1982

A/Title: The limited diversity of the mouse gamma-chains anti-GAT repertoire does not se

A/Reference number: S07453; MUID:83058021; PMID:6815271

A/Accession: S07453

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-43 <ROC2>

C/Superfamily: immunoglobulin V region; immunoglobulin homology

F:15-98/Domain: immunoglobulin homology <IMW>

Query Match 72.8%; Score 67; DB 2; Length 115;

Best Local Similarity 75.0%; Pred. No. 0.001;

Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFQ 16  
:|||||:|||||  
Db 50 RIDPANGTKYDPRFQ 65

## RESULT 10

S29594  
Ig gamma chain (WM65) - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999  
C/Accession: S29594  
R/Seymour, R.  
submitted to the EMBL Data Library, February 1991  
A/Reference number: S29593  
A/Accession: S29594  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-178 <SEV>  
A/Cross-references: EMBL:X57857; NID:g52590; PIDN:CAA40992.1; PID:g52591  
C/Keywords: immunoglobulin

Query Match 72.8%; Score 67; DB 2; Length 178;

Best Local Similarity 75.0%; Pred. No. 0.0016;

Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFQ 16  
:|||||:|||||  
Db 63 RIDPANGTKYDPRFQ 78

## RESULT 11

A27646  
Ig heavy chain V region (anti-haloperidol antibody E) - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 01-Dec-1989 #sequence\_revision 30-Sep-1991 #text\_change 21-Jan-2000  
C/Accession: A27646  
R/Sherman, M.A.; Deans, R.J.; Bolger, M.B.  
J. Biol. Chem. 263, 4059-4063, 1988  
A/Title: Haloperidol binding to monoclonal antibodies. Hypervariable region amino acid s

A/Reference number: A28195; MUID:88153717; PMID:3267217

A/Accession: A27646

A/Molecule type: mRNA

A/Residues: 1-107 <SHE>

A/Cross-references: GB:M17058

A/Note: the authors translated the codon TTC for residue 92 as Leu

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: heterotrimer; immunoglobulin

F:6-89/Domain: immunoglobulin homology <IMW>

Query Match 71.7%; Score 66; DB 2; Length 107;

Best Local Similarity 75.0%; Pred. No. 0.0013;

Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFQ 16  
:|||||:|||||  
Db 41 RIDPANGTKYDPRFQ 56

## RESULT 12

PH1013  
Ig heavy chain V region (clone 111.67) - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000  
C/Accession: PH1013  
R/Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.  
J. Exp. Med. 176, 761-779, 1992

A/Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B s

A/Reference number: PH0971; MUID:92381444; PMID:1512540

A/Accession: PH1013

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-107 <TIL>

C/Experimental source: B cell, strain [NZB x NZW]F1

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: heterotrimer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMW>

Query Match 70.7%; Score 65; DB 2; Length 107;  
Best Local Similarity 75.0%; Pred. No. 0.002;  
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16  
:||||| ||||| |||

Db 50 RIDPANGTKYAPKFQ 65

RESULT 13  
PH1012  
Ig heavy chain V region (clone 17p.73) - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000  
C/Accession: PH1012  
R/Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.  
J. Exp. Med. 176, 761-779, 1992  
A/Title: Both IGM and IGG anti-DNA antibodies are the products of clonally selective B cell  
A/Reference number: PH0971; MUID:92381444; PMID:1512540  
A/Accession: PH1012  
A/Status: nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 1-108 <TIL>  
A/Experimental source: B cell, strain [NZB x NZW]F1  
C/Superfamily: immunoglobulin V region; immunoglobulin homology  
C/Keywords: heterotetramer; immunoglobulin  
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 70.7%; Score 65; DB 2; Length 108;  
Best Local Similarity 75.0%; Pred. No. 0.002;  
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16  
:||||| ||||| |||

Db 50 RIDPANGTKYAPKFQ 65

RESULT 14  
A47271  
nitrophenyl phosphonate-specific antibody 48G7 heavy chain VDJ - synthetic (fragment)  
C/Species: synthetic  
A/Note: Mus musculus (house mouse) gene engineered and expressed in Escherichia coli  
C/Date: 21-Sep-1993 #sequence\_revision 11-Aug-1995 #text\_change 11-Aug-1995  
C/Accession: A47271  
R/Lealey, S.A.; Patten, P.A.; Schultz, P.G.  
Proc. Natl. Acad. Sci. U.S.A. 90, 1160-1165, 1993  
A/Title: A genetic approach to the generation of antibodies with enhanced catalytic acti  
A/Reference number: A47271; MUID:93165660; PMID:8094556  
A/Accession: A47271  
A/Molecule type: DNA; protein  
A/Residues: 1-114 <LES>  
A/Note: sequence extracted from NCBI backbone (NCBIN:124854, NCBIP:124855)  
A/Note: parts of this sequence were determined by protein sequencing  
F;22-96/Disulfide bonds: #status predicted

Query Match 68.5%; Score 63; DB 4; Length 114;  
Best Local Similarity 75.0%; Pred. No. 0.004;  
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16  
:||||| ||||| |||

Db 50 RIDPANGTKYDPIFQ 65

RESULT 15  
S24289  
Ig gamma chain V region (J534/32) - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Jun-2000  
C/Accession: S24289  
R/Monchamont, B.  
submitted to the EMBL Data Library, September 1991  
A/Description: Cloning and sequencing of the cDNA coding for the variable regions of the

A/Reference number: S24287  
A/Accession: S24289  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-116 <MON>  
A/Cross-references: EMBL:X62705; NID:G51690; PIDN:CAA44584.1; PID:gl333963  
C/Superfamily: immunoglobulin V region; immunoglobulin homology  
C/Keywords: immunoglobulin  
F;14-97/Domain: immunoglobulin homology <IMM>

Query Match 68.5%; Score 63; DB 2; Length 116;  
Best Local Similarity 68.8%; Pred. No. 0.0045;  
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16  
:||||| ||||| |||

Db 49 RIDPANGNTYDPIFQ 64

RESULT 16  
A56446  
Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)  
C/Species: Mus musculus (house mouse)  
C/Date: 19-Jan-1996 #sequence\_revision 19-Jan-1996 #text\_change 16-Aug-1996  
C/Accession: A56446  
R/Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.  
J. Biol. Chem. 270, 7829-7835, 1995  
A/Title: A high affinity digoxin-binding protein displayed on M13 is functionally identi  
A/Reference number: A56446; MUID:95229583; PMID:7713873  
A/Accession: A56446  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-268 <TAN>  
A/Cross-references: GB:U20617  
C/Keywords: heterotetramer; immunoglobulin

Query Match 67.4%; Score 62; DB 2; Length 268;  
Best Local Similarity 75.0%; Pred. No. 0.016;  
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFQ 16  
:||||| ||||| |||

Db 52 RIAPANGITKYDPIFQ 67

RESULT 17  
S52445  
Ig heavy chain V region precursor - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 08-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 21-Jan-2000  
C/Accession: S52445  
R/Berdoz, J.; Kraehenbuhl, J.P.  
submitted to the EMBL Data Library, November 1994  
A/Description: Specific amplification by the polymerase chain reaction of rearranged ge  
A/Reference number: S52445  
A/Accession: S52445  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-137 <BER>  
A/Cross-references: EMBL:X82690; NID:G673439; PIDN:CAA58011.1; PID:G673440  
C/Genetics:  
A/Introns: 16/1  
C/Superfamily: immunoglobulin V region; immunoglobulin homology  
C/Keywords: heterotetramer; immunoglobulin  
F;34-115/Domain: immunoglobulin homology <IMM>

Query Match 64.1%; Score 59; DB 2; Length 137;  
Best Local Similarity 73.3%; Pred. No. 0.024;  
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 IDPANGTKYDPIFQ 16  
||||| ||||| |||||

Db 70 IDPENGNTYDPIFQ 84

## RESULT 18

PL0246  
Ig heavy chain V region (anti-DNA, D23VH) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 21-Jan-2000  
C;Accession: PL0246  
R;Shlomchik, M.; Mascelli, M.; Shan, H.; Radic, M.Z.; Pisetsky, D.; Marshak-Rothstein, A.  
J. Exp. Med. 171, 263-297, 1990  
A;Title: Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic  
A;Reference number: PL0231; MUID:90111618; PMID:2104919  
A;Accession: PL0246  
A;Molecule type: mRNA  
A;Residues: 1-115 <SHL>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;1-29/Region: framework 1  
F;14-97/Domain: immunoglobulin homology <IMM>  
F;30-34/Region: complementarity-determining 1  
F;35-48/Region: framework 2  
F;43-65/Region: complementarity-determining 2  
F;66-97/Region: framework 3  
F;98-105/Region: complementarity-determining 3  
F;106-115/Region: framework 4

Query Match 60.9%; Score 56; DB 2; Length 115;  
Best Local Similarity 68.8%; Pred. No. 0.063;  
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFO 16  
:|||||:|||||  
Db 49 RIDSANGTKYVPK 64

## RESULT 19

S03484  
Ig heavy chain V-D-J region (hybridoma G7 Ab 2.9) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
A;Variety: strain BALB/c  
C;Date: 26-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 21-Jan-2000  
C;Accession: S03484; S07453  
R;Rocca-Serra, J.; Matthes, H.W.; Kaartinen, M.; Milstein, C.; Theze, J.; Fougereau, M.  
EMBO J. 2, 867-872, 1983  
A;Title: Analysis of antibody diversity: V-D-J mRNA nucleotide sequence of four anti-GAT  
hypervariable regions.  
A;Reference number: S03471; MUID:84057768; PMID:6416834  
A;Accession: S03484  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 10-120 <ROCI>  
A;Cross-references: EMBL:X07144  
A;Note: this sequence was determined from the differentiated gene  
R;Rocca-Serra, J.; Mazie, J.C.; Moinier, D.; Leclercq, L.; Somme, G.; Theze, J.; Fougere  
J. Immunol. 129, 2554-2558, 1982  
A;Title: The limited diversity of the mouse gamma-chains anti-GAT repertoire does not se  
A;Reference number: S07453; MUID:83058021; PMID:6815271  
A;Accession: S07453  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-43 <ROCI>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 59.8%; Score 55; DB 2; Length 120;  
Best Local Similarity 62.5%; Pred. No. 0.096;  
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFO 16  
:|||||:|||||  
Db 50 RIDPANKNSKYGP 65

## RESULT 20

S54495  
Probable carrier protein YPR021c - yeast (Saccharomyces cerevisiae)  
N;Alternate names: hypothetical protein YP9367.01c  
C;Species: Saccharomyces cerevisiae  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: S54495  
R;Badcock, K.; Churcher, C.M.  
Submitted to the EMBL Data Library, May 1995  
A;Reference number: S54059  
A;Accession: S54495  
A;Molecule type: DNA  
A;Residues: 1-902 <BAD>  
A;Cross-references: UNIPROT:Q12482; EMBL:Z49274; NID:g809585; PIDN:CAA89275.1; PID:g8095  
A;Experimental source: strain AB972  
C;Genetics:  
A;Gene: MIPS:YPR021c  
A;Cross-references: SGD:S0006225  
A;Map position: 16R  
C;Superfamily: probable carrier protein YPR021c; ADP,ATP carrier protein repeat homology  
C;Keywords: duplication; transmembrane protein  
F;527-615/Domain: ADP,ATP carrier protein repeat homology <ACP1>  
F;621-711/Domain: ADP,ATP carrier protein repeat homology <ACP2>  
F;724-814/Domain: ADP,ATP carrier protein repeat homology <ACP3>

Query Match 58.7%; Score 54; DB 1; Length 902;  
Best Local Similarity 58.8%; Pred. No. 1.3;  
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17  
:|||||:|||||  
Db 756 QIDPRKGETKNGIFHA 772

## RESULT 21

A36025  
Ig heavy chain V region (PR8-1) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 11-Jan-1991 #sequence\_revision 11-Jan-1991 #text\_change 23-Jul-1999  
C;Accession: A36025; E36025  
R;Caton, A.J.; Koprowski, H.  
Proc. Natl. Acad. Sci. U.S.A. 87, 6450-6454, 1990  
A;Title: Influenza virus hemagglutinin-specific antibodies isolated from a combinatorial  
A;Reference number: A36025; MUID:90349634; PMID:1696733  
A;Accession: A36025  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-82 <CAT>  
A;Cross-references: GB:M55999; NID:g194955; PIDN:AAA38112.1; PID:g194956  
A;Note: PR8-1  
A;Accession: E36025  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 19-82 <CA2>  
A;Cross-references: GB:M57273; NID:g194963; PIDN:AAA38116.1; PID:g194964  
A;Note: clone PR8-21  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin

Query Match 57.6%; Score 53; DB 2; Length 82;  
Best Local Similarity 66.7%; Pred. No. 0.13;  
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IDPANGTKYDPIFO 16  
:|||||:|||||  
Db 13 IDPENGDEYAPKFK 27

## RESULT 22

PH1144  
Ig heavy chain V region (clone V2048.2B) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004

C;Accession: PH1144  
R;Schitteck, B.; Rajewsky, K.  
J. Exp. Med. 176, 427-438, 1992  
A;Title: Natural occurrence and origin of somatically mutated memory B cells in mice.  
A;Reference number: PH1105; MUID:92364545; PMID:1500855  
A;Accession: PH1144  
A;Molecule type: DNA  
A;Residues: 1-98 <SCH>  
A;Cross-references: UNIPROT:Q924Q4; UNIPROT:Q924Q6; UNIPROT:Q924Q9; UNIPROT:Q924Q5; UNIPROT:Q924R0; UNIPROT:Q924R2  
A;Experimental source: B cell  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 53.3%; Score 49; DB 2; Length 98;  
Best Local Similarity 52.9%; Pred. No. 0.73;  
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17  
:||||:|||||  
Db 50 RIDPNSGTYDEKFKS 66

RESULT 23  
S54246  
Ig mu heavy chain V region precursor - sheep (fragment)  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 23-Jul-1999  
C;Accession: S54246  
R;Dufour, V.; Nau, F.  
submitted to the EMBL Data Library, April 1995  
A;Description: Sheep immunoglobulin mu heavy chain variable region sequence.  
A;Reference number: S54225  
A;Accession: S54246  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-138 <DUF>  
A;Cross-references: EMBL:Z49171; NID:g794133; PIDN:CAA89040.1; PID:g794134  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: immunoglobulin  
F;34-116/Domain: immunoglobulin homology <IMM>

Query Match 52.2%; Score 48; DB 2; Length 138;  
Best Local Similarity 66.7%; Pred. No. 1.6;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 6 NGTKYDPIFOA 17  
:||||:|||||  
Db 73 DGKTYDPAFQS 84

RESULT 24  
BMBYDL  
RAD50 protein - yeast (Saccharomyces cerevisiae)  
N;Alternate names: protein N0872; protein YNL250w  
C;Species: Saccharomyces cerevisiae  
C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C;Accession: S05808; S63223  
R;Alani, E.; Subbiah, S.; Kleckner, N.  
Genetics 122, 47-57, 1989  
A;Title: The yeast RAD50 gene encodes a predicted 153-kD protein containing a purine nuc  
A;Reference number: S05808; MUID:89276917; PMID:2659437  
A;Accession: S05808  
A;Molecule type: DNA  
A;Residues: 1-1312 <LA>  
A;Cross-references: UNIPROT:P12753; EMBL:X14814; NID:g4272; PIDN:CAA32919.1; PID:g4273  
R;Sen-Gupta, M.; Guedener, U.; Beinhauer, J.; Fiedler, T.; Hegemann, J.H.  
submitted to the Protein Sequence Database, April 1996  
A;Reference number: S63220  
A;Accession: S63223  
A;Molecule type: DNA  
A;Residues: 1-1312 <SEN>

C;Accession: PH1144  
R;Schitteck, B.; Rajewsky, K.  
J. Exp. Med. 176, 427-438, 1992  
A;Title: Natural occurrence and origin of somatically mutated memory B cells in mice.  
A;Reference number: PH1105; MUID:92364545; PMID:1500855  
A;Accession: PH1144  
A;Molecule type: DNA  
A;Residues: 1-98 <SCH>  
A;Cross-references: UNIPROT:Q924Q4; UNIPROT:Q924Q6; UNIPROT:Q924Q9; UNIPROT:Q924Q5; UNIPROT:Q924R0; UNIPROT:Q924R2  
A;Experimental source: B cell  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 53.3%; Score 49; DB 2; Length 98;  
Best Local Similarity 52.9%; Pred. No. 0.73;  
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17  
:||||:|||||  
Db 50 RIDPNSGTYDEKFKS 66

RESULT 23  
S54246  
Ig mu heavy chain V region precursor - sheep (fragment)  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 23-Jul-1999  
C;Accession: S54246  
R;Dufour, V.; Nau, F.  
submitted to the EMBL Data Library, April 1995  
A;Description: Sheep immunoglobulin mu heavy chain variable region sequence.  
A;Reference number: S54225  
A;Accession: S54246  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-138 <DUF>  
A;Cross-references: EMBL:Z49171; NID:g794133; PIDN:CAA89040.1; PID:g794134  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: immunoglobulin  
F;34-116/Domain: immunoglobulin homology <IMM>

Query Match 52.2%; Score 48; DB 2; Length 138;  
Best Local Similarity 66.7%; Pred. No. 1.6;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 6 NGTKYDPIFOA 17  
:||||:|||||  
Db 73 DGKTYDPAFQS 84

RESULT 24  
BMBYDL  
RAD50 protein - yeast (Saccharomyces cerevisiae)  
N;Alternate names: protein N0872; protein YNL250w  
C;Species: Saccharomyces cerevisiae  
C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C;Accession: S05808; S63223  
R;Alani, E.; Subbiah, S.; Kleckner, N.  
Genetics 122, 47-57, 1989  
A;Title: The yeast RAD50 gene encodes a predicted 153-kD protein containing a purine nuc  
A;Reference number: S05808; MUID:89276917; PMID:2659437  
A;Accession: S05808  
A;Molecule type: DNA  
A;Residues: 1-1312 <LA>  
A;Cross-references: UNIPROT:P12753; EMBL:X14814; NID:g4272; PIDN:CAA32919.1; PID:g4273  
R;Sen-Gupta, M.; Guedener, U.; Beinhauer, J.; Fiedler, T.; Hegemann, J.H.  
submitted to the Protein Sequence Database, April 1996  
A;Reference number: S63220  
A;Accession: S63223  
A;Molecule type: DNA  
A;Residues: 1-1312 <SEN>

A;Cross-references: EMBL:Z71526; NID:g1302292; PIDN:CAA96157.1; PID:g1302293; GSPDB:GN001  
A;Experimental source: strain S288C  
C;Genetics:  
A;Gene: SGD:RAD50; MIPS:YNL250w  
A;Cross-references: SGD:S0005194; MIPS:YNL250w  
A;Map position: 14L  
C;Superfamily: RAD50 protein  
C;Keywords: ATP; coiled coil; DNA repair; meiosis; nucleus  
F;177-421/Region: heptad repeats  
F;743-995/Region: heptad repeats  
F;40/Binding site: ATP (lys) #status predicted

Query Match 52.2%; Score 48; DB 1; Length 1312;  
Best Local Similarity 60.0%; Pred. No. 18;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 DPANGTKYDPIFOA 17  
:||||:|||||  
Db 167 EPSNLKKKDFEIFOA 181

RESULT 25  
PH1429  
Ig heavy chain V region (clone VH5-2R1) - human (fragment)  
N;Alternate names: Ig heavy chain V-III region (TD-Vn)  
C;Species: Homo sapiens (man)  
C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 21-Jan-2000  
C;Accession: PH1429; PLO119  
R;van der Stoep, N.; van der Linden, J.; Logtenberg, T.  
J. Exp. Med. 177, 99-107, 1993  
A;Title: Molecular evolution of the human immunoglobulin E response: High incidence of si  
dermatitis.  
A;Reference number: PH1409; MUID:93115676; PMID:8418213  
A;Accession: PH1429  
A;Molecule type: mRNA  
A;Residues: 1-98 <VAN>  
R;Bird, J.; Galili, N.; Link, M.; Stites, D.; Sklar, J.  
J. Exp. Med. 168, 229-245, 1988  
A;Title: Continuing rearrangement but absence of somatic hypermutation in immunoglobulin  
A;Reference number: PLO116; MUID:88286083; PMID:2840480  
A;Accession: PLO119  
A;Molecule type: mRNA  
A;Residues: 1-98 <BIR>  
A;Experimental source: B cells from patient TD with acute lymphoblastic leukemia, ALL  
C;Note: the sequence shows the V region (TD-Vn) from one of five DNA rearrangements from  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: acute lymphoblastic leukemia; heterotetramer; immunoglobulin  
F;15-98/Domain: immunoglobulin homology <IMM>  
F;31-35/Region: complementarity-determining 1  
F;49-65/Region: complementarity-determining 2

Query Match 51.1%; Score 47; DB 2; Length 98;  
Best Local Similarity 50.0%; Pred. No. 1.6;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFO 16  
:||||:|||||  
Db 50 RIDPDSYTYNSPSFQ 65

RESULT 26  
S12431  
Ig heavy chain V region (clone 5RG) - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 12-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 21-Jan-2000  
C;Accession: S12431  
R;Sanz, I.; Kelly, P.; Williams, C.; Scholl, S.; Tucker, P.; Capra, J.D.  
EMBO J. 8, 3741-3748, 1989  
A;Title: The smaller human V(H) gene families display remarkably little polymorphism.  
A;Reference number: S09421; MUID:90059975; PMID:2511001  
A;Accession: S12431  
A;Status: translation not shown  
A;Molecule type: DNA



A;Residues: 1-101 <SRN>  
A;Cross-references: EMBL:X56374  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
F;18-101/Domain: immunoglobulin homology <IMM>

Query Match 51.1%; Score 47; DB 2; Length 101;  
Best Local Similarity 50.0%; Pred. No. 1.6; Mismatches 3; Indels 0; Gaps 0;  
Matches 8; Conservative 3;

Qy 1 KIDPANGKTKYDPIFQ 16  
:||||: |||||  
Db 53 RIDPSDSYTNYSFSFQ 68

## RESULT 27

PH1560

Ig heavy chain V region (clone VH32) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 05-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 21-Jan-2000

C;Accession: PH1560

R;Rassenti, L.Z.; Kipps, T.J.

J. Exp. Med. 177, 1039-1046, 1993

A;Title: Lack of extensive mutations in the VH5 genes used in common B cell chronic lymph

A;Reference number: PH1557; MUID:93210459; PMID:7681468

A;Accession: PH1560

A;Molecule type: DNA

A;Residues: 1-115 &lt;RAS&gt;

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;32-115/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 51.1%; Score 47; DB 2; Length 115;  
Best Local Similarity 50.0%; Pred. No. 1.9; Mismatches 3; Indels 0; Gaps 0;  
Matches 8; Conservative 3;

Qy 1 KIDPANGKTKYDPIFQ 16  
:||||: |||||  
Db 67 RIDPSDSYTNYSFSFQ 82

## RESULT 28

PH1426

Ig heavy chain V region (clone P2-57) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 21-Jan-2000

C;Accession: PH1426

R;van der Stoep, N.; van der Linden, J.; Logtenberg, T.

J. Exp. Med. 177, 99-107, 1993

A;Title: Molecular evolution of the human immunoglobulin B response: High incidence of dermatitis.

A;Reference number: PH1409; MUID:93115676; PMID:8418213

A;Accession: PH1426

A;Molecule type: mRNA

A;Residues: 1-122 &lt;VAN&gt;

A;Experimental source: PBMC

A;Note: the authors translated the codon AAA for residue 28 as Ser and GTA for residue 8

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-98/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 51.1%; Score 47; DB 2; Length 122;  
Best Local Similarity 50.0%; Pred. No. 2; Mismatches 3; Indels 0; Gaps 0;  
Matches 8; Conservative 3;

Qy 1 KIDPANGKTKYDPIFQ 16  
:||||: |||||  
Db 50 RIDPSDSYTNYSFSFQ 65

## RESULT 29

PH1562

Ig heavy chain V region (clone HOW) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 05-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 21-Jan-2000

C;Accession: PH1562

R;Rassenti, L.Z.; Kipps, T.J.

J. Exp. Med. 177, 1039-1046, 1993

A;Title: Lack of extensive mutations in the VH5 genes used in common B cell chronic lymph

A;Reference number: PH1557; MUID:93210459; PMID:7681468

A;Accession: PH1562

A;Molecule type: DNA

A;Residues: 1-137 &lt;RAS&gt;

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;32-115/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 51.1%; Score 47; DB 2; Length 137;  
Best Local Similarity 50.0%; Pred. No. 2.2; Mismatches 3; Indels 0; Gaps 0;  
Matches 8; Conservative 3;

Qy 1 KIDPANGKTKYDPIFQ 16  
:||||: |||||  
Db 67 RIDPSDSYTNYSFSFQ 82

## RESULT 30

PH1564

Ig heavy chain V region (clone ANG) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 05-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 21-Jan-2000

C;Accession: PH1564

R;Rassenti, L.Z.; Kipps, T.J.

J. Exp. Med. 177, 1039-1046, 1993

A;Title: Lack of extensive mutations in the VH5 genes used in common B cell chronic lymph

A;Reference number: PH1557; MUID:93210459; PMID:7681468

A;Accession: PH1564

A;Molecule type: DNA

A;Residues: 1-138 &lt;RAS&gt;

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;32-115/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 51.1%; Score 47; DB 2; Length 138;  
Best Local Similarity 50.0%; Pred. No. 2.3; Mismatches 3; Indels 0; Gaps 0;  
Matches 8; Conservative 3;

Qy 1 KIDPANGKTKYDPIFQ 16  
:||||: |||||  
Db 67 RIDPSDSYTNYSFSFQ 82

## RESULT 31

PH1565

Ig heavy chain V region (clone KER) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 05-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 21-Jan-2000

C;Accession: PH1565

R;Rassenti, L.Z.; Kipps, T.J.

J. Exp. Med. 177, 1039-1046, 1993

A;Title: Lack of extensive mutations in the VH5 genes used in common B cell chronic lymph

A;Reference number: PH1557; MUID:93210459; PMID:7681468

A;Accession: PH1565

A;Molecule type: DNA

A;Residues: 1-138 &lt;RAS&gt;

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;32-115/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 51.1%; Score 47; DB 2; Length 138;  
Best Local Similarity 50.0%; Pred. No. 2.3; Mismatches 3; Indels 0; Gaps 0;  
Matches 8; Conservative 3;

Qy 1 KIDPANGKTKYDPIFQ 16  
:||||: |||||  
Db 67 RIDPSDSYTNYSFSFQ 82

```
RESULT 32
PH1563
Ig heavy chain V region (clone PET) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 05-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 21-Jan-2000
C;Accession: PH1563
R;Rassenti, L.Z.; Kipps, T.J.
J. Exp. Med. 177, 1039-1046, 1993
A;Title: Lack of extensive mutations in the VHS genes used in common B cell chronic lymphocytic leukemia
A;Reference number: PH1557; UID:93210459; PMID:7681468
A;Accession: PH1563
A;Molecule type: DNA
A;Residues: 1-144 <RAS>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotrimer; immunoglobulin
F;32-115/Domain: immunoglobulin homology <IMM>

Query Match      51.1%; Score 47; DB 2; Length 144;
Best Local Similarity 50.0%; Pred. No. 2.4;
Matches      8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      1 KIDPANGKTKYDPIFQ 16
Db      67 RIDPDSYTNYSFSFQ 82

RESULT 33
PH1561
Ig heavy chain V region (clone CAV) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 05-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 21-Jan-2000
C;Accession: PH1561
R;Rassenti, L.Z.; Kipps, T.J.
J. Exp. Med. 177, 1039-1046, 1993
A;Title: Lack of extensive mutations in the VHS genes used in common B cell chronic lymphocytic leukemia
A;Reference number: PH1557; UID:93210459; PMID:7681468
A;Accession: PH1561
A;Molecule type: DNA
A;Residues: 1-147 <RAS>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotrimer; immunoglobulin
F;32-115/Domain: immunoglobulin homology <IMM>

Query Match      51.1%; Score 47; DB 2; Length 147;
Best Local Similarity 50.0%; Pred. No. 2.4;
Matches      8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      1 KIDPANGKTKYDPIFQ 16
Db      67 RIDPDSYTNYSFSFQ 82

RESULT 34
PH1427
Ig heavy chain V region (clone P1-58) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jan-2000
C;Accession: PH1427
R;van der Stoep, N.; van der Linden, J.; Logtenberg, T.
J. Exp. Med. 177, 99-107, 1993
A;Title: Molecular evolution of the human immunoglobulin E response: High incidence of dermatitis.
A;Reference number: PH1409; UID:93115676; PMID:8418213
A;Accession: PH1427
A;Molecule type: mRNA
A;Residues: 1-132 <VAN>
A;Experimental source: PBMC
A;Note: the authors translated the codon TCT for residue 57 as His, ACT for residue 58 as Thr, and TGT for residue 59 as Cys.
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotrimer; immunoglobulin
F;15-98/Domain: immunoglobulin homology <IMM>
```

```
Query Match      50.0%; Score 46; DB 2; Length 132;
Best Local Similarity 50.0%; Pred. No. 3.1;
Matches      8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      1 KIDPANGKTKYDPIFQ 16
Db      50 RIDPDSSTNYCPSFQ 65

RESULT 35
S41374
single chain Fv antibody - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C;Accession: S41374
R;Artsenko, O.; Weiler, E.W.; Muentz, K.; Conrad, U.
submitted to the EMBL Data Library, January 1994
A;Description: Construction and functional characterization of a single chain Fv antibody
A;Reference number: S41374
A;Accession: S41374
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <ART>
A;Cross-references: EMBL:Z29480

Query Match      50.0%; Score 46; DB 2; Length 249;
Best Local Similarity 56.2%; Pred. No. 6.3;
Matches      9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy      1 KIDPANGKTKYDPIFQ 16
Db      50 RIAPASGLTKYVPRFQ 65

RESULT 36
H69186
conserved hypothetical protein MTH653 - Methanobacterium thermoautotrophicum (strain Del
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: H69186
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.;
; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
A;Reference number: A69000; UID:98037514; PMID:9371463
A;Accession: H69186
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-507 <MTH>
A;Cross-references: UNIPROT:O26749; GB:AE000845; GB:AE000666; NID:g2621726; PIDN:AAB85151;
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH653

Query Match      50.0%; Score 46; DB 2; Length 507;
Best Local Similarity 52.9%; Pred. No. 14;
Matches      9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      1 KIDPANGKTKYDPIFOA 17
Db      44 KVTMANGLTFFDPIYNA 60

RESULT 37
AC0238
hematin storage system, HmsF protein [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0238
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
; M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.
```

Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AC0238  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-673 <KUR>  
A:Cross-references: UNIPROT:Q9R7V4; GB:AL590842; PIDN:CAC90767.1; PID:gl5979969; GSPDB:C  
C:Genetics:  
A:Gene: hmsF  
C:Superfamily: Yersinia pestis hypothetical protein hmsF

Query Match 50.0%; Score 46; DB 2; Length 673;  
Best Local Similarity 61.5%; Pred. No. 19;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDIP 13  
:|||||  
Db 417 RIDPKTGKTSIDP 429

## RESULT 38

T47006  
hypothetical protein hmsF [imported] - Yersinia pestis  
C:Species: Yersinia pestis  
C:Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: T47006  
R:Buchrieser, C.; Runiok, C.; Couve, E.; Frangeul, L.; Billault, A.; Kunst, F.; Carniel  
submitted to the EMBL Data Library, October 1998  
A:Description: DNA sequence of the 102 kbases unstable region of Yersinia pestis.  
A:Reference number: 224348

A:Accession: T47006  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-673 <BUC>  
A:Cross-references: UNIPROT:Q9R7V4; EMBL:AL031866; PIDN:CAA21349.1  
A:Experimental source: strain 6/69  
C:Genetics:  
A:Note: hmsF

C:Superfamily: Yersinia pestis hypothetical protein hmsF

Query Match 50.0%; Score 46; DB 2; Length 673;  
Best Local Similarity 61.5%; Pred. No. 19;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDIP 13  
:|||||  
Db 417 RIDPKTGKTSIDP 429

## RESULT 39

PH1139  
Ig heavy chain V region (clone V2042.2B) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C:Accession: PH1139  
R:Schitteck, B.; Rajewsky, K.  
J. Exp. Med. 176, 427-438, 1992  
A:Title: Natural occurrence and origin of somatically mutated memory B cells in mice.  
A:Reference number: PH1105; MUID:92364545; PMID:1500855

A:Accession: PH1139  
A:Molecule type: DNA  
A:Residues: 1-98 <SCH>  
A:Cross-references: UNIPROT:Q924Q4; UNIPROT:Q924Q6; UNIPROT:Q924Q9; UNIPROT:Q924Q5; UNIPROT:Q924R4; UNIPROT:Q924R0; UNIPROT:Q924Q8; UNIPROT:Q924R2; UNIPROT:Q9K172  
A:Experimental source: B cell  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 48.9%; Score 45; DB 2; Length 98;  
Best Local Similarity 47.1%; Pred. No. 3.3;  
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDIPFOA 17  
:|||||  
Db 50 RIDPKSGGTYNKFKS 66

## RESULT 40

S25174  
Ig heavy chain V region - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 21-Jan-2000  
C:Accession: S25174; S31133  
R:Monestier, M.; Fasy, T.M.; Losman, M.J.; Novick, K.E.; Muller, S.  
submitted to the EMBL Data Library, July 1992  
A:Description: Structure and binding properties of monoclonal antibodies to core histone  
A:Reference number: S25174  
A:Accession: S25174  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-118 <MON>

A:Cross-references: EMBL:X67622; NID:g51854; PIDN:CAA47880.1; PID:g938259  
A:Accession: S31133  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-118 <MO2>  
A:Cross-references: EMBL:X67622; NID:g51854; PIDN:CAA47880.1; PID:g938259  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 48.9%; Score 45; DB 2; Length 118;  
Best Local Similarity 60.0%; Pred. No. 4;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 IDPANGTKYDIPFQ 16  
:|||||  
Db 51 IDPENGDTYASKFQ 65

Search completed: June 3, 2005, 16:02:32  
Job time : 29.8788 secs

This Page Blank (upside)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 15:40:50 ; Search time 144.242 Seconds  
(without alignments)  
60.352 Million cell updates/sec

Title: US-10-089-452-22

Perfect score: 92

Sequence: 1 KIDPANGKTKYDPIFQA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_crembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	62	67.4	109	2 Q9JL85	Q9JL85 mus musculus
2	54	58.7	902	2 Q12482	Q12482 saccharomyc
3	48	52.2	1296	2 Q752D9	Q752D9 ashbya goss
4	48	52.2	1305	2 Q6FNZ8	Q6FNZ8 candida gla
5	48	52.2	1312	1 RA50 YEAST	P12753 saccharomyc
6	47	51.1	633	2 Q8LG54	Q8LG54 oryza sativ
7	47	51.1	643	2 Q8CX16	Q8CX16 oceanobacil
8	46	50.0	241	2 Q6MW59	Q6MW59 neurospora
9	46	50.0	507	2 Q26749	Q26749 methanobact
10	46	50.0	673	2 Q56940	Q56940 yersinia ps
11	46	50.0	673	2 Q66B30	Q66B30 yersinia pe
12	46	50.0	673	2 Q9R7V4	Q9R7V4 yersinia pe
13	45	48.9	206	2 Q8L968	Q8L968 arabidopsis
14	45	48.9	231	2 Q9SVF0	Q9SVF0 arabidopsis
15	45	48.9	933	2 Q6BML1	Q6BML1 debaryomyce
16	45	48.9	1337	2 Q9ZC4	Q9ZC4 mus musculus
17	45	48.9	1500	1 GRLF CANFA	P83509 canis famil
18	45	48.9	1513	1 GRLF HUMAN	Q9NRY4 homo sapien
19	45	48.9	1513	1 GRLF RAT	P81128 rattus norv
20	44	47.8	117	1 HV09 MOUSE	P01753 mus musculus
21	44	47.8	117	1 HV09 MOUSE	P01754 mus musculus
22	44	47.8	137	2 Q924R6	Q924R6 mus musculus
23	44	47.8	139	1 HV07 MOUSE	P01751 mus musculus
24	44	47.8	139	2 Q924R5	Q924R5 mus musculus
25	44	47.8	140	2 Q924R2	Q924R2 mus musculus
26	44	47.8	141	2 Q924Q4	Q924Q4 mus musculus
27	44	47.8	143	2 Q91VA2	Q91VA2 mus musculus
28	44	47.8	143	2 Q924Q5	Q924Q5 mus musculus
29	44	47.8	143	2 Q924R0	Q924R0 mus musculus
30	44	47.8	143	2 Q924R7	Q924R7 mus musculus
31	44	47.8	144	2 Q924P5	Q924P5 mus musculus

32	44	47.8	145	2 Q924P7	Q924P7 mus musculus
33	44	47.8	145	2 Q924Q6	Q924Q6 mus musculus
34	44	47.8	145	2 Q924Q7	Q924Q7 mus musculus
35	44	47.8	145	2 Q924Q9	Q924Q9 mus musculus
36	44	47.8	145	2 Q924R1	Q924R1 mus musculus
37	44	47.8	145	2 Q924R3	Q924R3 mus musculus
38	44	47.8	145	2 Q924R4	Q924R4 mus musculus
39	44	47.8	146	2 Q924Q3	Q924Q3 mus musculus
40	44	47.8	146	2 Q924Q8	Q924Q8 mus musculus
41	44	47.8	146	2 Q924R8	Q924R8 mus musculus
42	44	47.8	267	2 Q8T222	Q8T222 pyrococcus
43	44	47.8	356	2 Q6MQ10	Q6MQ10 bdellovibri
44	44	47.8	388	2 Q87Y17	Q87Y17 pseudomonas
45	44	47.8	460	1 PPAN_DROME	Q9YDES drosophila

#### ALIGNMENTS

##### RESULT 1

Q9JL85 PRELIMINARY; PRT; 109 AA.  
AC Q9JL85  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Anti-myosin immunoglobulin heavy chain variable region  
DE (fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c;  
RX MEDLINE=20448942; PubMed=10992488;  
RX DOI=10.1128/JAI.68.10.5803-5808.2000;  
RA Malkiel S., Liao L., Cunningham M.W., Diamond B.;  
RT "T-Cell-dependent antibody response to the dominant epitope of  
RT streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive  
RT with cardiac myosin."  
RL Infect. Immun. 68:5803-5808(2000).  
DR EMBL; AF206021; AAP69319.1; -.  
DR HSBP; P01751; INQB.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003596; Ig\_v.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS50835; IGV LIKE; 1.  
FT NON\_TER 1  
FT NON\_TER 109  
SQ SEQUENCE 109 AA; 11944 MW; DFE615FB6CED4EDE CRC64;

Query Match 67.4%; Score 62; DB 2; Length 109;

Best Local Similarity 68.8%; Pred. No. 0.028;

Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFQ 16

Db 42 RIDPATGHSKYDPKFK 57

##### RESULT 2

Q12482 PRELIMINARY; PRT; 902 AA.

AC Q12482;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein YPR020C.

GN Name=AGC1; Synonyms=YPR020C;

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

```

OX NCBI_TaxID=4932;
RP SEQUENCE FROM N.A.
RC STRAIN=AB972;
RA Badcock K., Bowman S., Churcher C.M., Pearson D., Rajandream M.A.,
RA Walsh S.V., Barrell B.G.;
RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RL
[2]
RP SEQUENCE FROM N.A.
RP STRAIN=AB972;
RC Badcock K., Churcher C.M.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=AB972;
RA Barrell B., Rajandream M.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the mitochondrial carrier family.
DR EMBL; Z71255; CAA95017.1; -.
DR EMBL; Z49274; CAA89275.1; -.
DR PIR; S54495; S54495.
DR SGD; S000006225; AGC1.
DR GO; GO:0015183; P:L-aspartate transporter activity; IDA.
DR GO; GO:0005313; P:L-glutamate transporter activity; IDA.
DR GO; GO:0015810; P:L-aspartate transporter; IDA.
DR GO; GO:0015813; P:L-glutamate transporter; IDA.
DR InterPro; IPR002113; Adenine transporter.
DR InterPro; IPR001993; Mitochondrial carrier.
DR Pfam; PF00153; Mitochondrial carrier.
DR PRINTS; PR00927; ADPTRNSLCASE.
DR PRINTS; PR00926; MITOCARRIER.
DR PROSITE; PS00920; SOLCAR; 3.
KW Hypothetical protein; Transmembrane; Transporter.
SQ SEQUENCE 902 AA; 104303 MW; 07D6F831E2CD15CF CRC64;

Query Match 58.7%; Score 54; DB 2; Length 902;
Best Local Similarity 58.8%; Pred. No. 5.8;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17
Db 756 QIDPRKGETKINGIFHA 772
:|||||:|||||:

RESULT 3
ID Q752D9 PRELIMINARY; PRT; 1296 AA.
AC Q752D9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AFR637WP.
GN ORFNames=AFR637W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RA Brachat S., Voegelé S.E., Dietrich F.S., Lerch A., Gaffney T.,
RA Philippsen P.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016819; AAS54008.1; -.
DR AGD; AFR637W; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0030870; C:Mre11 complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity; coupled to transmembrane m. . .; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; ABC_transporter.

Query Match 52.2%; Score 48; DB 2; Length 1305;
Best Local Similarity 60.0%; Pred. No. 86;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 DPANGTKYDPIFOA 17
Db 167 EPSNLKKKFDPIFOA 181
:|||||:|||||:

```

```

DR InterPro; IPR004584; Rad50.
DR InterPro; IPR007517; Rad50 zn_hook.
DR InterPro; IPR003395; SMC_N.
DR Pfam; PF04423; Rad50 zn_hook; 1.
DR Pfam; PF02463; SMC_N; 1.
DR TIGRFAMs; TIGR00606; rad50; 1.
SQ SEQUENCE 1296 AA; 149398 MW; 74987F28CDF37251 CRC64;

Query Match 52.2%; Score 48; DB 2; Length 1296;
Best Local Similarity 60.0%; Pred. No. 86;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 DPANGTKYDPIFOA 17
Db 166 EPANLKKKFDPIFOA 180
:|||||:|||||:

RESULT 4
ID Q6FNZ8 PRELIMINARY; PRT; 1305 AA.
AC Q6FNZ8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to sp|P12753 Saccharomyces cerevisiae YNL250w RAD50 DNA repair protein.
GN ORFNames=CAGL0J077889;
OS Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=284593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller R.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
EMBL; CR380956; CAG60997.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0030870; C:Mre11 complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity; coupled to transmembrane m. . .; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR004584; Rad50.
DR InterPro; IPR007517; Rad50 zn_hook.
DR Pfam; PF04423; Rad50 zn_hook; 1.
DR TIGRFAMs; TIGR00606; rad50; 1.
SQ SEQUENCE 1305 AA; 150667 MW; 74898660D317CEF80 CRC64;

Query Match 52.2%; Score 48; DB 2; Length 1305;
Best Local Similarity 60.0%; Pred. No. 86;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 DPANGTKYDPIFOA 17
Db 167 EPSNLKKKFDPIFOA 181
:|||||:|||||:

```

Query Match 52.2%; Score 48; DB 1; Length 1312;  
Best Local Similarity 60.0%; Pred. No. 87;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 DPANGTKYVDPIFOA 17  
DB 167 EFSNLKXKXFDPIFOA 181

RESULT 6  
Q9LGS9 PRELIMINARY; PRT; 633 AA.

ID Q9LGS9  
AC Q9LGS9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Similar to Arabidopsis thaliana chromosome I BAC F25P22.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzoae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=12447438; DOI=10.1038/nature01184;  
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
RA Ikono M., Itoh S., Itoh Y., Itoh Y., Iwabuchi N., Kono I.,  
RA Karasawa M., Katagiri S., Kikuta A., Kobayashi N., Komiya K.,  
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,  
RA Shinokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
RA Zhong H., Iwana H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
RA Yano M., Jiang J., Gojobori T.;  
RT "The genome sequence and structure of rice chromosome 1.";  
RL Nature 420:312-316(2002).  
DR EMBL; AF002538; BAB03388.1; --  
DR Gramene; Q9LGS94; --  
SQ SEQUENCE 633 AA; 71207 MW; 3DCE9D1C01E41C22 CRC64;

Query Match 51.1%; Score 47; DB 2; Length 633;  
Best Local Similarity 61.5%; Pred. No. 57;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDIP 13  
DB 613 KVDPAQGWQDRDP 625

RESULT 7  
Q8CXI6 PRELIMINARY; PRT; 643 AA.

ID Q8CXI6  
AC Q8CXI6;  
DT 01-WAR-2003 (TrEMBLrel. 23, Created)  
DT 01-WAR-2004 (TrEMBLrel. 23, Last sequence update)  
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Stage V sporulation protein D (Sporulation specific penicillin-binding  
DE protein) (Spore cortex).  
GN Name=spovD; OrderedLocustNames=OBI465;  
OS Oceanobacillus theysensis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.  
OX NCBI\_TaxID=182710;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HT5831;  
RX MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkf526;  
RA Takami H., Takaki Y., Uchiyama I.;  
RT "Genome sequence of Oceanobacillus theysensis isolated from the Iheya

Query Match 52.2%; Score 48; DB 1; Length 1312;  
Best Local Similarity 60.0%; Pred. No. 87;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 DPANGTKYVDPIFOA 17  
DB 167 EFSNLKXKXFDPIFOA 181

RESULT 6  
Q9LGS9 PRELIMINARY; PRT; 633 AA.

ID Q9LGS9  
AC Q9LGS9;  
DT 01-OCT-2000 (TREMELrel. 15, Created)  
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)  
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)  
DE Similar to Arabidopsis thaliana chromosome I BAC F25P22.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaeae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=12447438; DOI=10.1038/nature01184;  
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
RA Ikono M., Itoh S., Itoh Y., Itoh Y., Iwabuchi N., Kono I.,  
RA Karasawa M., Katagiri S., Kikuta A., Kobayashi N., Komiya K.,  
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,  
RA Shinokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
RA Zhong H., Iwana H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
RA Yano M., Jiang J., Gojobori T.;  
RT "The genome sequence and structure of rice chromosome 1.";  
RL Nature 420:312-316(2002).  
DR EMBL; AF002538; BAB03388.1; --  
DR Gramene; Q9LGS94; --  
SQ SEQUENCE 633 AA; 71207 MW; 3DCE9D1C01E41C22 CRC64;

Query Match 51.1%; Score 47; DB 2; Length 633;  
Best Local Similarity 61.5%; Pred. No. 57;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KIDPANGTKYKYP 13  
DB 613 KVDPAQGMQRDP 625

RESULT 7  
Q8CXI6 PRELIMINARY; PRT; 643 AA.

ID Q8CXI6  
AC Q8CXI6;  
DT 01-WAR-2003 (TREMELrel. 23, Created)  
DT 01-WAR-2003 (TREMELrel. 23, Last sequence update)  
DT 01-WAR-2004 (TREMELrel. 26, Last annotation update)  
DE Stage V sporulation protein D (Sporulation specific penicillin-binding  
DE protein) (Spore cortex).  
GN Name=spoVD; OrderedLocustNames=OBI465;  
OS Oceanobacillus theysensis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.  
OX NCBI\_TaxID=182710;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HT5831;  
RX MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkf526;  
RA Takami H., Takaki Y., Uchiyama I.;  
RT "Genome sequence of Oceanobacillus theysensis isolated from the Iheya

```

RT Ridge and its unexpected adaptive capabilities to extreme
RL environments."
DR Nucleic Acids Res. 30:3927-3935 (2002).
DR EMBL; AP004598; BAC13421.1; -.
DR HSSP; O34006; 1K25.
DR GO; GO:0008658; P:penicillin binding; IEA.
DR CO; GO:0009273; P:cell wall biosynthesis (sensu Bacteria); IEA.
DR InterPro; IPR005543; PASTA.
DR InterPro; IPR005311; PBP dimer.
DR InterPro; IPR001460; Pencil_bind_tpept.
DR Pfam; PF03793; PASTA; 1.
DR Pfam; PF03717; PBP dimer; 1.
DR Pfam; PF00905; Transpeptidase; 1.
KW Complete proteome.
SQ SEQUENCE 643 AA; 69936 MW; 78C080A52F4BFF35 CRC64;

Query Match 51.1%; Score 47; DB 2; Length 643;
Best Local Similarity 58.8%; Pred. No. 58;
Matches 10; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

Qy 2 IDPANGKT--KYDPIFQ 16
Db 443 IDPVNGETVEKYEQLQ 459

RESULT 8
Q6MW59
ID Q6MW59 PRELIMINARY; PRT; 241 AA.
AC Q6MW59;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein B4B2.050.
GN Name=B4B2.050;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hobeisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RA Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RA Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; BX842615; CAE76030.1; -.
KW Hypothetical protein.
SQ SEQUENCE 241 AA; 25347 MW; 616ED3FAD117FF38 CRC64;

Query Match 50.0%; Score 46; DB 2; Length 241;
Best Local Similarity 90.0%; Pred. No. 29;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ANGKTKYDPI 14
Db 91 ANGKTKTDPI 100

RESULT 9
ID O26749 PRELIMINARY; PRT; 507 AA.
AC O26749;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Conserved protein.
GN OrderedLocustNames=MTF653;
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzaden R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155 (1997).
DR EMBL; AE000845; AAB85158.1; -.
DR PIR; H69186; H69186.
KW Complete proteome.
SQ SEQUENCE 507 AA; 57324 MW; 02E2A56ADF90CF17 CRC64;

Query Match 50.0%; Score 46; DB 2; Length 507;
Best Local Similarity 52.9%; Pred. No. 66;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFOA 17
Db 44 KYVMANGLTFDFPIYNA 60

RESULT 10
Q56940
ID Q56940 PRELIMINARY; PRT; 673 AA.
AC Q56940;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE HmsF.
GN Name=hmsF;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM6+;
RX MEDLINE=97390123; PubMed=9249062; DOI=10.1016/S0378-1119(97)00071-1;
RA Lillard J.W. Jr., Fetherston J.D., Pedersen L., Pendrak M.L.,
RA Perry R.D.;
RT "Sequence and genetic analysis of the hemin storage (hms) system of
RT Yersinia pestis.";
RL Gene 193:13-21 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM6+;
RX Lillard J.W.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM6+;
RA Perry R.D.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; U22837; AAB6589.2; -.
DR GO; GO:0016810; P:hydrolyase activity, acting on carbon-nitrog. .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR002509; Polysac deacet.
DR Pfam; PF01522; Polysacc deac_1; 1.
SQ SEQUENCE 673 AA; 76685 MW; 39F1F9E7C52D7D46 CRC64;

Query Match 50.0%; Score 46; DB 2; Length 673;
Best Local Similarity 61.5%; Pred. No. 90;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDP 13
Db 44 KYVMANGLTFDFPIYNA 60

```



```
Db 417 RIDPKTGKTSIDP 429

RESULT 11
Q66B30 PRELIMINARY; PRT; 673 AA.
AC Q66B30;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Haemin storage system, HmsF protein precursor.
GN Name=hmsF; ORFNames=YPT1950;
OS Versinia pseudotuberculosis IP 32953.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=273123;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP 32953;
RX Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin L.V.,
RA Brubaker R.R., Fowler J., Hinnebusch B.J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the genome evolution of Versinia pestis through whole
RT genome comparison with Versinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH21188.1; -
DR InterPro; IPR002509; Polysac_deacet.
DR Pfam; PF01522; Polysacc_deac_1; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 673 AA; 76713 MW; 19D1876C3A66AF02 CRC64;

Query Match 50.0%; Score 46; DB 2; Length 673;
Best Local Similarity 61.5%; Pred. No. 90;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDIP 13
:|||||
Db 417 RIDPKTGKTSIDP 429

RESULT 12
Q9R7V4 PRELIMINARY; PRT; 673 AA.
AC Q9R7V4; Q7AWD0; Q7CI65;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Haemin storage system, HmsF protein (Hypothetical protein Y2358).
GN Name=hmsF; OrderedLocNames=Yp1696, Yp01952, Y2358;
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=6/69;
RA Buchrieser C., Rusniok C., Couve E., Frangeul L., Billault A.,
RA Kunst P., Carniel E., Glaser P.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;

RT "Genome sequence of Versinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Churcher C.M., Mungall K.L.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.P., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Versinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RX Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031866; CAA21349.1; -
DR EMBL; AE013838; AAM85916.1; -
DR EMBL; AJ414150; CAC90767.1; -
DR EMBL; AE017133; AAS61926.1; -
DR PIR; AC0238; AC0238.
DR PIR; T47006; T47006.
DR GO; GO:0016810; P:hydrolase activity, acting on carbon-nitrog. .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR002509; Polysac_deacet.
DR Pfam; PF01522; Polysacc_deac_1; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 673 AA; 76713 MW; 19D1876C3A66AF02 CRC64;

Query Match 50.0%; Score 46; DB 2; Length 673;
Best Local Similarity 61.5%; Pred. No. 90;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDIP 13
:|||||
Db 417 RIDPKTGKTSIDP 429

RESULT 13
Q8L968 PRELIMINARY; PRT; 206 AA.
AC Q8L968;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative HAM1 protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=22088475; PubMed=12093376;
RX Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN RP SEQUENCE FROM N.A.
RC STRAIN=V.V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the HAM1 NTPase family.
CC EMBL; AY088611; AAM66934.1; -
```

DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR InterPro; IPR002637; Hamlp\_like.  
 DR Pfam; PF01725; Hamlp\_like; 1.  
 DR TIGRFAMs; TIGR00042; Hamlp\_like; 1.  
 KW Hydrolase.  
 SQ SEQUENCE 206 AA; 22478 MW; C86C23082A57CAB1 CRC64;  
 Query Match 48.9%; Score 45; DB 2; Length 206;  
 Best Local Similarity 55.6%; Pred. No. 36;  
 Matches 10; Conservative 2; Mismatches 4; Indels 2; Gaps 1;  
 Qy 1 KIDPANGTKY--DPFQ 16  
 ||||| : |||||  
 Db 142 KIVPARGPTDFGDPVFQ 159  
 ||||| : |||||  
 RESULT 14  
 ID Q9SVF0 PRELIMINARY; PRT; 231 AA.  
 AC Q9SVF0;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein F18A5.110 (Hypothetical protein AT4g13720).  
 GN Name=F18A5.110; Synonyms=AT4g13720;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Weber N., Grueninger D., Schmidheini T., Bancroft I.,  
 RA Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBSJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBSJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Weber N., Grueninger D., Schmidheini T., Mewes H.W., Lemcke K.,  
 RA Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.  
 DR EMBL; AL035528; CAB36836.1; -;  
 DR EMBL; AL161537; CAB78414.1; -;  
 DR PIR; T05241; T05241.  
 DR HSSP; Q57679; 1B78.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR InterPro; IPR002637; Hamlp\_like.  
 DR Pfam; PF01725; Hamlp\_like; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 231 AA; 25382 MW; A2480804572C4220 CRC64;  
 Query Match 48.9%; Score 45; DB 2; Length 231;  
 Best Local Similarity 55.6%; Pred. No. 41;  
 Matches 10; Conservative 2; Mismatches 4; Indels 2; Gaps 1;  
 Qy 1 KIDPANGTKY--DPFQ 16  
 ||||| : |||||  
 Db 167 KIVPARGPTDFGDPVFQ 184  
 ||||| : |||||  
 RESULT 15  
 ID Q6BML1 PRELIMINARY; PRT; 913 AA.  
 AC Q6BML1;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Similar to CA5878|CaPRP5 Candida albicans CaPRP5 pre-mRNA processing  
 DE RNA-helicase.  
 GN ORFNames=DEHA0F04895g;  
 OS Debaryomyces hansenii CBS767.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.  
 OX NCBI\_TaxID=284592;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RG Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boisrame A., Boyer J., Catrhead C., Confanioli F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Swemene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts";  
 RL Nature 430:35-44(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RG Genoscope;  
 RA Submitted (JUL-2004) to the EMBL/GenBank/DBSJ databases.  
 RL -1- SIMILARITY: Belongs to the DEAD box helicase family.  
 CC EMBL; CR382138; CAG88878.1; -;  
 DR GO; GO:0005224; F:ATP binding; IEA.  
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR011545; DEAD/DEAH\_N.  
 DR InterPro; IPR000629; DEAD\_box\_N.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; Helicase\_C; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELICG; 1.  
 DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
 DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
 KW ATP-binding; Helicase; Hydrolase.  
 SQ SEQUENCE 913 AA; 102639 MW; 8B59782049005F8D CRC64;  
 Query Match 48.9%; Score 45; DB 2; Length 913;  
 Best Local Similarity 64.3%; Pred. No. 1.8e+02;  
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 Qy 2 IDPANGTKYDPIF 15  
 ||||| : |||||  
 Db 140 ITPANGELKKRP 153  
 ||||| : |||||  
 RESULT 16  
 ID Q69ZC4 PRELIMINARY; PRT; 1337 AA.  
 AC Q69ZC4;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE MKIAA1722 protein (Fragment).  
 GN Name=MKIAA1722;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Soga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RA "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:205-218(2004).
DR EMBL; AK173242; BAD32520.1; -
DR InterPro; IPR002713; FF.
DR InterPro; IPR00198; RhoGAP.
DR InterPro; IPR008936; Rho_GAP.
DR Pfam; PF01846; FF; 2.
DR Pfam; PF00620; RhoGAP; 1.
DR SMART; SM00441; FF; 4.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS0238; RHO GAP; 1.
DR NON TER 1
FT SEQUENCE 1337 AA; 152050 MW; 3E5511675813C77D CRC64;
SQ
Query Match 48.9%; Score 45; DB 2; Length 1337;
Best Local Similarity 56.2%; Pred. No. 2.8e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 IDPANGTKYDPIFOA 17
  ||:|||||:|
Db 85 IDKSRGKTIIPYFEA 100

RESULT 17
GRLF CANFA
ID _GRLF CANFA STANDARD; PRT; 1500 AA.
AC P83509;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Glucocorticoid receptor DNA binding factor 1 (Rho GAP p190A) (p190-A).
GN Name=GRFL1;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=22220127; PubMed=12234678; DOI=10.1016/S0378-1119(02)00765-5;
RA Zangerl B., Zhang Q., Pearce-Kelling S.E., Aguirre G.D.;
RT "Molecular cloning, characterization and mapping of the canine
RT glucocorticoid receptor DNA binding factor 1 (GRFL1).";
RL Gene 294:167-176(2002).
CC -1- FUNCTION: Represses transcription of the glucocorticoid receptor
CC by binding to the cis-acting regulatory sequence 5'-
CC GAGAAAGAACTGGAGAACTC-3'. May participate in the regulation of
CC retinal development and degeneration. May transduce signals from
CC p21-ras to the nucleus, acting via the ras GTP-ase activating
CC protein (GAP). May also act as a tumor suppressor (By similarity).
CC -1- SUBUNIT: Interacts with p120GAP (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: Strongly expressed in retina (photoreceptor
CC layer) and brain. Expression is maximal in the occipital, frontal,
CC temporal lobe and also the cerebellum. Medium expression in the
CC medulla and also in kidney, lung, liver, heart and spleen.
CC -1- PTM: Tyrosine phosphorylated (By similarity).
CC -1- SIMILARITY: Contains 4 FF domains.
CC -1- SIMILARITY: Contains 1 Rho-GAP domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL; AF483595; AAN16354.1; -
DR EMBL; AY079157; AAL91068.1; -
DR EMBL; AY079158; AAL91069.1; -
DR HSP; Q98935; I77C.
DR GO; GO:0005634; C:nucleus; ISS.
DR GO; GO:0016564; F:transcriptional repressor activity; ISS.
DR GO; GO:0016481; P:negative regulation of transcription; ISS.
DR GO; GO:0000074; P:regulation of cell cycle; ISS.
DR GO; GO:0042478; P:regulation of eye photoreceptor cell developme. . .; NAS.
DR InterPro; IPR002713; FF.
DR InterPro; IPR001806; Ras_trnsfrmg.
DR InterPro; IPR008936; Rho_GAP.
DR InterPro; IPR00198; RhoGAP.
DR Pfam; PF01846; FF; 2.
DR Pfam; PF00620; RhoGAP; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00441; FF; 4.
DR SMART; SM00324; RhoGAP; 1.
DR PROSITE; PS0238; RHO GAP; 1.
KW Anti-oncogene; DNA-binding; GTPase activation; Nuclear protein;
KW Phosphorylation; Repeat; Repressor; Transcription regulation.
FT DOMAIN 270 327 FF 1.
FT DOMAIN 369 422 FF 2.
FT DOMAIN 429 483 FF 3.
FT DOMAIN 485 539 FF 4.
FT DOMAIN 1250 1437 Rho-GAP.
FT DOMAIN 1441 1485 Pro-rich.
SQ SEQUENCE 1500 AA; 170426 MW; EB3AB65FE36E2F18 CRC64;

Query Match 48.9%; Score 45; DB 1; Length 1500;
Best Local Similarity 56.2%; Pred. No. 3.1e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 IDPANGTKYDPIFOA 17
  ||:|||||:|
Db 247 IDKSRGKTIIPYFEA 262

RESULT 18
GRLF HUMAN
ID _GRLF HUMAN STANDARD; PRT; 1513 AA.
AC Q9NRV4; Q14452; Q9COE1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Glucocorticoid receptor DNA binding factor 1 (Glucocorticoid receptor
DE repression factor 1) (GRF-1) (Rho GAP p190A) (p190-A).
GN Name=GRFL1; Synonyms=GRF1, KIAA1722;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Breast cancer;
RX MEDLINE=20510021; PubMed=11054565; DOI=10.1016/S0378-1119(00)00387-5;
RA Tikoo A., Czekay S., Viars C., White S., Heath J.K., Arden K.,
RA Maruta H.;
RT "p190-A, a human tumor suppressor gene, maps to the chromosomal region
RT 19q13.3 that is reportedly deleted in some gliomas.";
RL Gene 257:23-31(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohata O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code

```

for large proteins in vitro.";  
 RL DNA Res. 7:347-355(2000).  
 [3]  
 RN SEQUENCE OF 377-1453 FROM N.A. (ISOFORM 1), AND FUNCTION.  
 RC TISSUE=Breast cancer; PubMed=1894621;  
 RX MEDLINE=91373352;  
 RA Leclerc S., Palaniwami R., Xie B.X., Govindan M.V.;  
 RT "Molecular cloning and characterization of a factor that binds the  
 human glucocorticoid receptor gene and represses its expression.";  
 J. Biol. Chem. 266:17333-17340(1991).  
 CC -1- FUNCTION: Represses transcription of the glucocorticoid receptor  
 by binding to the cis-acting regulatory sequence 5'-  
 GAGAAAGAACTGGAGAACTC-3'. May participate in the regulation of  
 retinal development and degeneration. May transduce signals from  
 p21-ras to the nucleus, acting via the ras GTP-ase activating  
 protein (GAP). May also act as a tumor suppressor.  
 CC -1- SUBUNIT: Interacts with p120GAP.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q9NRY4-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9NRY4-2; Sequence=VSP\_007105;  
 CC Note=No experimental confirmation available;  
 CC -1- PTM: Tyrosine phosphorylated (By similarity).  
 CC -1- SIMILARITY: Contains 4 FF domains.  
 CC -1- SIMILARITY: Contains 1 Rho-GAP domain.  
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to frameshifts  
 in positions 533, 540, 607 and 614.  
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to frameshifts  
 in positions 389, 533, 540, 607, 614, 1167, 1241, 1292, 1334 and  
 1446.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF159851; AAF80386.1; ALT\_FRAME.  
 DR EMBL; AB051509; BAB21813.2; -.  
 DR EMBL; M73077; AAA58618.1; ALT\_FRAME.  
 DR HSP; Q98935; 1F7C.  
 DR TRANSFAC; T00940; -.  
 DR Genew; HGNC:4591; GRLF1.  
 DR MIM; 605277; -.  
 DR GO; GO:0005634; C:nucleus; IC.  
 DR GO; GO:0016564; P:transcriptional repressor activity; IDA.  
 DR GO; GO:0016481; P:negative regulation of transcription; IDA.  
 DR GO; GO:0000074; P:regulation of cell cycle; NAS.  
 DR GO; GO:0042478; P:regulation of eye photoreceptor cell develop. . .; ISS.  
 DR InterPro; IPR002713; FF.  
 DR InterPro; IPR001806; Ras trnsmfrmg.  
 DR InterPro; IPR008936; Rho GAP.  
 DR InterPro; IPR00198; RhoGAP.  
 DR Pfam; PF01846; FF; 2.  
 DR Pfam; PF00620; RhoGAP; 1.  
 DR PRINTS; PR00449; RASTRNSFRMG.  
 DR SMART; SM00441; FF; 4.  
 DR SMART; SM00324; RhoGAP; 1.  
 DR PROSITE; PS50238; RHOAP; 1.  
 DR Alternative splicing; Anti-oncogene; DNA-binding; GTPase activation;  
 KW Nuclear protein; Phosphorylation; Repeat; Repressor;  
 KW Transcription regulation.  
 FT DOMAIN 270 327 FF 1.  
 FT DOMAIN 327 369 FF 2.  
 FT DOMAIN 369 422 FF 3.  
 FT DOMAIN 429 483 FF 4.  
 FT DOMAIN 485 539 FF 4.  
 FT DOMAIN 1249 1436 Rho-GAP.  
 FT DOMAIN 1440 1507 Pro-rich.

FT VARSPLIC 1492 1513 PLPHRPETPOPSRNPWQTKHEH -> QLQAEHTL (in  
 isoform 2).  
 FT /FTID=VSP\_007105.  
 FT R -> P (in Ref. 1 and 3).  
 FT V -> D (in Ref. 1 and 3).  
 FT S -> G (in Ref. 1 and 3).  
 FT W -> M (in Ref. 3).  
 FT Q -> A (in Ref. 1 and 3).  
 FT M -> T (in Ref. 1 and 3).  
 FT C -> S (in Ref. 1).  
 FT M -> I (in Ref. 1).  
 FT PS -> RN (in Ref. 3).  
 SQ SEQUENCE 1513 AA; 172227 MW; 305C2B3D06519F96 CRC64;  
 Query Match 48.9%; Score 45; DB 1; Length 1513;  
 Best Local Similarity 56.2%; Pred. No. 3.2e+02;  
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 IDPANGTKYDPIFQA 17  
 || : ||| | : |  
 Db 247 IDKSRGKTKIIPYFEA 262

RESULT 19  
 GRLF RAT  
 ID GRLF RAT STANDARD; PRT; 1513 AA.  
 AC P81128;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE Glucocorticoid receptor DNA binding factor 1 (GAP-associated protein  
 p190).  
 DE [1]  
 GN Name=Gr1f1;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., SEQUENCE OF 379-391; 491-506 AND 1304-1322, AND  
 PHOSPHORYLATION  
 RX MEDLINE=92257594; PubMed=1581965; DOI=10.1016/0092-8674(92)90454-K;  
 RA Settleman J., Narasiman V., Foster L.C., Weinberg R.A.;  
 RT "Molecular cloning of cDNAs encoding the GAP-associated protein p190:  
 implications for a signaling pathway from ras to the nucleus.";  
 Cell 69:539-549(1992).  
 CC -1- FUNCTION: Represses transcription of the glucocorticoid receptor  
 by binding to the cis-acting regulatory sequence 5'-  
 GAGAAAGAACTGGAGAACTC-3'. May transduce signals from p21-ras to  
 the nucleus, acting via the ras GTP-ase activating protein (GAP).  
 CC May also act as a tumor suppressor (By similarity).  
 CC -1- SUBUNIT: Interacts with p120GAP (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic, when bound to GAP and also  
 partially found to be associated with nuclei.  
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.  
 CC -1- PTM: Tyrosine phosphorylated.  
 CC -1- SIMILARITY: Contains 4 FF domains.  
 CC -1- SIMILARITY: Contains 1 Rho-GAP domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M94721; -. NOT\_ANNOTATED\_CDS.  
 DR PIR; A38218; A38218.  
 DR HSP; Q98935; 1F7C.  
 DR InterPro; IPR002713; FF.  
 DR InterPro; IPR001806; Ras trnsmfrmg.  
 DR InterPro; IPR008936; Rho GAP.  
 DR InterPro; IPR00198; RhoGAP.

```

DR Pfam; PF01846; PF; 2.
DR Pfam; PF00071; Rab; 1.
DR Pfam; PF00620; RHOGAP; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00441; PF; 4.
DR SMART; SM00324; RHOGAP; 1.
DR PROSITE; PS0238; RHOGAP; 1.
KW Anti-oncogene; Direct protein sequencing; DNA-binding;
KW GTPase activation; Nuclear protein; Phosphorylation; Repeat;
KW Repressor; Transcription regulation.
FT DOMAIN 270 327 FF 1.
FT DOMAIN 369 422 FF 2.
FT DOMAIN 429 483 FF 3.
FT DOMAIN 485 539 FF 4.
FT DOMAIN 1249 1436 Rho-GAP.
SQ SEQUENCE 1513 AA; 172378 MW; D99D02C036EB5519 CRC64;

Query Match 48.9%; Score 45; DB 1; Length 1513;
Best Local Similarity 56.2%; Pred. No. 3.2e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 IDPANGTKYDPIFOA 17
Db 247 IDKSRGKTKIIPYFEA 262

RESULT 20
HV09 MOUSE
ID HV09 MOUSE STANDARD; PRT; 117 AA.
AC P01753; P11271;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 186-1 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376; DOI=10.1016/0092-8674(81)90089-1;
RA Botwell A.L.M., Paskind M., Reth M., Inanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the Npb family of
antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -1- MISCELLANEOUS: This germline gene belongs to a set of closely
related genes that could encode V regions of Npb antibodies.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J00533; AAA38602.1; -.
DR PIR; C90809; HVMS45.
DR HSSP; P01751; 1A6W.
DR MGD; MGI:96486; IGH-VJ558.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 Ig heavy chain V region 145.
FT DOMAIN 20 49 Framework-1.
FT DOMAIN 50 54 Complementarity-determining-1.
FT DOMAIN 55 68 Framework-2.
FT DOMAIN 69 85 Complementarity-determining-2.
FT DOMAIN 86 117 Framework-3.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12921 MW; D37D8A3F543E996 CRC64;

Query Match 47.8%; Score 44; DB 1; Length 117;
Best Local Similarity 47.1%; Pred. No. 28;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17
Db 69 RIDPNSGGTKYNERFKS 85

RESULT 22
Q924R6
ID Q924R6 PRELIMINARY; PRT; 137 AA.
AC Q924R6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

```

```
DE VH186.2-D-J-C mu protein (Fragment).
GN Name=VH186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB067783; BAB63268.1; -
DR PIR: F28833; F28833.
DR PIR: PH1105; PH1105.
DR PIR: PH1108; PH1108.
DR PIR: PH1112; PH1112.
DR PIR: PH1114; PH1114.
DR PIR: PH1118; PH1118.
DR PIR: PH1119; PH1119.
DR PIR: PH1122; PH1122.
DR PIR: PH1123; PH1123.
DR PIR: PH1124; PH1124.
DR PIR: PH1125; PH1125.
DR PIR: PH1126; PH1126.
DR PIR: PH1128; PH1128.
DR PIR: PH1129; PH1129.
DR PIR: PH1131; PH1131.
DR PIR: PH1134; PH1134.
DR PIR: PH1137; PH1137.
DR PIR: PH1139; PH1139.
DR PIR: PH1142; PH1142.
DR PIR: PH1144; PH1144.
DR PIR: PH1145; PH1145.
DR PIR: PH1147; PH1147.
DR PIR: PH1149; PH1149.
DR PIR: PH1150; PH1150.
DR PIR: PH1151; PH1151.
DR PIR: PH1152; PH1152.
DR PIR: PH1153; PH1153.
DR HSSP: P01751; IA6W.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 137
FT NON_TER 137
SQ SEQUENCE 137 AA; 15171 MW; 5C38D966DC6A4124 CRC64;

Query Match 47.8%; Score 44; DB 2; Length 137;
Best Local Similarity 47.1%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Oy 1 KIDPANGTKYDPIFOA 17
Db 50 RIDPNSGGTKYNEKFKS 66
:||||:||||:|::

RESULT 23
HV07 MOUSE STANDARD; PRT; 139 AA.
AC P01751; P01752;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ig heavy chain V region B1-8/186-2 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA MEDLINE=81234548; PubMed=6788376; DOI=10.1016/0092-8674(81)90089-1;
RA Bothwell A.L.M., Taskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
```

```
RT "Heavy chain variable region contribution to the NPb family of
RL antibodies: somatic mutation evident in a gamma 2a variable region.";
Cell 24:625-637(1981).
CC -I- MISCELLANEOUS: The B1-8 mu chain mRNA was cloned from a hybridoma
CC making antibodies to the haptan (4-hydroxy-3-nitrophenyl)acetyl
CC (NPb antibodies).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00529; AAA38170.1; -.
DR PIR: A90809; MEMS18.
DR PDB; 1A6U; X-ray; H=20-139.
DR PDB; 1A6V; X-ray; H/I/J=20-139.
DR PDB; 1A6W; X-ray; H=20-139.
DR PDB; 1NOB; X-ray; A/C=20-139.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR 3D-structure; Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 139 Ig heavy chain V region B1-8/186-2.
FT DOMAIN 20 49 Framework-1.
FT DOMAIN 50 54 Complementarity-determining-1.
FT DOMAIN 55 68 Framework-2.
FT DOMAIN 69 85 Complementarity-determining-2.
FT DOMAIN 86 117 Framework-3.
FT DOMAIN 118 124 D segment.
FT DOMAIN 125 139 JH2 segment.
FT DISULFID 41 115 By similarity.
FT STRAND 22 24
FT STRAND 28 31
FT STRAND 33 34
FT STRAND 37 44
FT STRAND 48 50
FT HELIX 52 58
FT TURN 60 61
FT STRAND 64 70
FT TURN 72 74
FT STRAND 77 79
FT HELIX 81 83
FT TURN 84 86
FT STRAND 87 92
FT TURN 93 96
FT STRAND 97 102
FT HELIX 107 109
FT STRAND 111 118
FT TURN 120 123
FT STRAND 129 129
FT STRAND 133 137
FT NON_TER 139 139
SQ SEQUENCE 139 AA; 15419 MW; 1B57DD4FD0C9F465 CRC64;

Query Match 47.8%; Score 44; DB 1; Length 139;
Best Local Similarity 47.1%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Oy 1 KIDPANGTKYDPIFOA 17
Db 69 RIDPNSGGTKYNEKFKS 85
:||||:||||:|::

RESULT 24
Q924R5 PRELIMINARY; PRT; 139 AA.
ID Q924R5
AC Q924R5;
```

```

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHI186.2-D-J-C mu protein (fragment).
GN Name=VHI186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067784; BAB63283.1; -.
DR PIR; PH1137; PH1137.
DR HSP; P01751; IAGW.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 139
SQ SEQUENCE 139 AA; 15221 MW; 8491E2F85614736A CRC64;

Query Match 47.8%; Score 44; DB 2; Length 139;
Best Local Similarity 47.1%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFOA 17
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 25
Q924R2 PRELIMINARY; PRT; 140 AA.
AC Q924R2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHI186.2-D-J-C mu protein (fragment).
GN Name=VHI186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067788; BAB63273.1; -.
DR PIR; F28833; F28833.
DR PIR; PH1105; PH1105.
DR PIR; PH1108; PH1108.
DR PIR; PH1114; PH1114.
DR PIR; PH1118; PH1118.
DR PIR; PH1119; PH1119.
DR PIR; PH1125; PH1125.
DR PIR; PH1126; PH1126.
DR PIR; PH1128; PH1128.
DR PIR; PH1129; PH1129.
DR PIR; PH1131; PH1131.
DR PIR; PH1134; PH1134.
DR PIR; PH1137; PH1137.
DR PIR; PH1139; PH1139.
DR PIR; PH1142; PH1142.
DR PIR; PH1144; PH1144.
DR PIR; PH1147; PH1147.
DR PIR; PH1149; PH1149.
DR PIR; PH1150; PH1150.
DR PIR; PH1151; PH1151.
DR PIR; PH1152; PH1152.
DR PIR; PH1153; PH1153.

Query Match 47.8%; Score 44; DB 2; Length 139;
Best Local Similarity 47.1%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFOA 17
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 26
Q924Q4 PRELIMINARY; PRT; 141 AA.
AC Q924Q4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHI186.2-D-J-C mu protein (fragment).
GN Name=VHI186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067796; BAB63281.1; -.
DR PIR; F28833; F28833.
DR PIR; PH1105; PH1105.
DR PIR; PH1108; PH1108.
DR PIR; PH1114; PH1114.
DR PIR; PH1118; PH1118.
DR PIR; PH1119; PH1119.
DR PIR; PH1125; PH1125.
DR PIR; PH1126; PH1126.
DR PIR; PH1128; PH1128.
DR PIR; PH1129; PH1129.
DR PIR; PH1131; PH1131.
DR PIR; PH1134; PH1134.
DR PIR; PH1137; PH1137.
DR PIR; PH1139; PH1139.
DR PIR; PH1142; PH1142.
DR PIR; PH1144; PH1144.
DR PIR; PH1147; PH1147.
DR PIR; PH1149; PH1149.
DR PIR; PH1150; PH1150.
DR PIR; PH1151; PH1151.
DR PIR; PH1152; PH1152.
DR PIR; PH1153; PH1153.
DR HSP; P01751; IAGW.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 141
SQ SEQUENCE 141 AA; 15561 MW; DDD80482D66B76A0 CRC64;

Query Match 47.8%; Score 44; DB 2; Length 141;
Best Local Similarity 47.1%; Pred. No. 35;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFOA 17
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 27
Q924Q4 PRELIMINARY; PRT; 141 AA.
AC Q924Q4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHI186.2-D-J-C mu protein (fragment).
GN Name=VHI186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067796; BAB63281.1; -.
DR PIR; F28833; F28833.
DR PIR; PH1105; PH1105.
DR PIR; PH1108; PH1108.
DR PIR; PH1114; PH1114.
DR PIR; PH1118; PH1118.
DR PIR; PH1119; PH1119.
DR PIR; PH1125; PH1125.
DR PIR; PH1126; PH1126.
DR PIR; PH1128; PH1128.
DR PIR; PH1129; PH1129.
DR PIR; PH1131; PH1131.
DR PIR; PH1134; PH1134.
DR PIR; PH1137; PH1137.
DR PIR; PH1139; PH1139.
DR PIR; PH1142; PH1142.
DR PIR; PH1144; PH1144.
DR PIR; PH1147; PH1147.
DR PIR; PH1149; PH1149.
DR PIR; PH1150; PH1150.
DR PIR; PH1151; PH1151.
DR PIR; PH1152; PH1152.
DR PIR; PH1153; PH1153.
DR HSP; P01751; IAGW.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 141
SQ SEQUENCE 141 AA; 15561 MW; DDD80482D66B76A0 CRC64;

Query Match 47.8%; Score 44; DB 2; Length 141;
Best Local Similarity 47.1%; Pred. No. 35;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFOA 17
Db 50 RIDPNSGGTKYNEKFKS 66
```

```

RESULT 27
Q91VA2 ID Q91VA2 PRELIMINARY; PRT; 143 AA.
AC Q91VA2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE VHI186.2-D-J-C mu protein (Fragment).
GN Name=VHI186.2-D-J-C mu;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067786; BAB63271.1; -.
DR EMBL; AB069911; BAB63927.1; -.
DR PIR; PH1105; PH1105.
DR PIR; PH1114; PH1114.
DR PIR; PH1118; PH1118.
DR PIR; PH1119; PH1119.
DR PIR; PH1125; PH1125.
DR PIR; PH1126; PH1126.
DR PIR; PH1128; PH1128.
DR PIR; PH1134; PH1134.
DR PIR; PH1137; PH1137.
DR PIR; PH1142; PH1142.
DR PIR; PH1149; PH1149.
DR PIR; PH1150; PH1150.
DR PIR; PH1151; PH1151.
DR PIR; PH1152; PH1152.
DR HSP; P01751; IAGW.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 143
SQ SEQUENCE 143 AA; 15617 MW; 51952152F6F3AD47 CRC64;

Query Match 47.8%; Score 44; DB 2; Length 143;
Best Local Similarity 47.1%; Pred. No. 35;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFOA 17
Db :|||:||||:|::
50 RIDPNSGGTKYNEKFKS 66

RESULT 28
Q924Q5 ID Q924Q5 PRELIMINARY; PRT; 143 AA.
AC Q924Q5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHI186.2-D-J-C mu protein (Fragment).
GN Name=VHI186.2-D-J-C mu;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067795; BAB63280.1; -.
DR PIR; F28833; F28833.
DR PIR; PH1105; PH1105.
DR PIR; PH1108; PH1108.
DR PIR; PH1114; PH1114.

Qy 1 KIDPANGKTKYDPIFOA 17
Db :|||:||||:|::
50 RIDPNSGGTKYNEKFKS 66

RESULT 29
Q924R0 ID Q924R0 PRELIMINARY; PRT; 143 AA.
AC Q924R0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHI186.2-D-J-C mu protein (Fragment).
GN Name=VHI186.2-D-J-C mu;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067790; BAB63275.1; -.
DR PIR; F28833; F28833.
DR PIR; F39932; F39932.
DR PIR; PH1105; PH1105.
DR PIR; PH1108; PH1108.
DR PIR; PH1114; PH1114.
DR PIR; PH1118; PH1118.
DR PIR; PH1119; PH1119.
DR PIR; PH1125; PH1125.
DR PIR; PH1126; PH1126.
DR PIR; PH1128; PH1128.
DR PIR; PH1129; PH1129.
DR PIR; PH1131; PH1131.
DR PIR; PH1134; PH1134.
DR PIR; PH1137; PH1137.
DR PIR; PH1139; PH1139.
DR PIR; PH1142; PH1142.
DR PIR; PH1144; PH1144.
DR PIR; PH1147; PH1147.
DR PIR; PH1149; PH1149.
DR PIR; PH1150; PH1150.

Qy 1 KIDPANGKTKYDPIFOA 17
Db :|||:||||:|::
50 RIDPNSGGTKYNEKFKS 66

```



```
DR PIR; PH1151; PH1151...
DR PIR; PH1152; PH1152...
DR PIR; PH1153; PH1153...
DR HSSP; P01751; 1A6W.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 143
FT NON_TER 143
SQ SEQUENCE 143 AA; 15868 MW; 139B2E966B81E07F CRC64;

Query Match 47.8%; Score 44; DB 2; Length 143;
Best Local Similarity 47.1%; Pred. No. 35;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFOA 17
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 30
Q924R7 PRELIMINARY; PRT; 143 AA.
AC Q924R7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHI86.2-D-J-C mu protein (Fragment).
GN Name=VHI86.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067782; BAB63267.1; -.
DR PIR; PH1105; PH1105.
DR PIR; PH1108; PH1108.
DR PIR; PH1114; PH1114.
DR PIR; PH1118; PH1118.
DR PIR; PH1119; PH1119.
DR PIR; PH1125; PH1125.
DR PIR; PH1126; PH1126.
DR PIR; PH1128; PH1128.
DR PIR; PH1131; PH1131.
DR PIR; PH1134; PH1134.
DR PIR; PH1137; PH1137.
DR PIR; PH1139; PH1139.
DR PIR; PH1142; PH1142.
DR PIR; PH1149; PH1149.
DR PIR; PH1150; PH1150.
DR PIR; PH1151; PH1151.
DR PIR; PH1152; PH1152.
DR PIR; PH1153; PH1153.
DR HSSP; P01751; 1A6W.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 143
FT NON_TER 143
SQ SEQUENCE 143 AA; 15648 MW; 51894D22EA9FDD47 CRC64;

Query Match 47.8%; Score 44; DB 2; Length 143;
Best Local Similarity 47.1%; Pred. No. 35;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFOA 17
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 31
Q924P5 PRELIMINARY; PRT; 144 AA.
AC Q924P5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHI86.2-D-J-C mu protein (Fragment).
GN Name=VHI86.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB069920; BAB63936.1; -.
DR PIR; PH1137; PH1137.
DR PIR; S26744; S26744.
DR HSSP; P01751; 1A6W.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON_TER 144
FT NON_TER 144
SQ SEQUENCE 144 AA; 15865 MW; CD07BE97E95C1B27 CRC64;

Query Match 47.8%; Score 44; DB 2; Length 144;
Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGKTKYDPIFOA 17
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 32
Q924P7 PRELIMINARY; PRT; 145 AA.
AC Q924P7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHI86.2-D-J-C mu protein (Fragment).
GN Name=VHI86.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB069918; BAB63934.1; -.
DR PIR; F28833; F28833.
DR PIR; F33932; F33932.
DR PIR; F33932; F33932.
DR PIR; PH1105; PH1105.
DR PIR; PH1114; PH1114.
DR PIR; PH1118; PH1118.
DR PIR; PH1119; PH1119.
DR PIR; PH1125; PH1125.
DR PIR; PH1126; PH1126.
DR PIR; PH1128; PH1128.
DR PIR; PH1134; PH1134.
DR PIR; PH1137; PH1137.
DR PIR; PH1142; PH1142.
DR PIR; PH1149; PH1149.
DR PIR; PH1150; PH1150.
DR PIR; PH1151; PH1151.
DR PIR; PH1152; PH1152.
DR PIR; PH1153; PH1153.
DR HSSP; P01751; 1A6W.
```

```

DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON TER 1
FT NON TER 145
SQ SEQUENCE 145 AA; 15988 MW; 73958704796C9A CRC64;

Query Match 47.8%; Score 44; DB 2; Length 145;
Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFOA 17
:||||:||||:|:|:
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 33
Q924Q6 PRELIMINARY; PRT; 145 AA.
AC Q924Q6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHL186.2-D-J-C mu protein (Fragment).
GN Names=VHL186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR PIR; F28833; F28833.
DR PIR; F28833; F28833.
DR PIR; PH1105; PH1105.
DR PIR; PH1108; PH1108.
DR PIR; PH1114; PH1114.
DR PIR; PH1118; PH1118.
DR PIR; PH1119; PH1119.
DR PIR; PH1125; PH1125.
DR PIR; PH1126; PH1126.
DR PIR; PH1128; PH1128.
DR PIR; PH1129; PH1129.
DR PIR; PH1131; PH1131.
DR PIR; PH1134; PH1134.
DR PIR; PH1137; PH1137.
DR PIR; PH1139; PH1139.
DR PIR; PH1142; PH1142.
DR PIR; PH1144; PH1144.
DR PIR; PH1147; PH1147.
DR PIR; PH1149; PH1149.
DR PIR; PH1150; PH1150.
DR PIR; PH1151; PH1151.
DR PIR; PH1152; PH1152.
DR PIR; PH1153; PH1153.
DR HSSP; P01751; 1A6W.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON TER 1
FT NON TER 145
SQ SEQUENCE 145 AA; 16141 MW; 55A59A7908B2CD6A CRC64;

Query Match 47.8%; Score 44; DB 2; Length 145;
Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFOA 17
:||||:||||:|:|:
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 35
Q924Q9 PRELIMINARY; PRT; 145 AA.
AC Q924Q9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHL186.2-D-J-C mu protein (Fragment).
GN Names=VHL186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;

```

```

DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON TER 1
FT NON TER 145
SQ SEQUENCE 145 AA; 15988 MW; 73958704796C9A CRC64;

Query Match 47.8%; Score 44; DB 2; Length 145;
Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFOA 17
:||||:||||:|:|:
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 33
Q924Q6 PRELIMINARY; PRT; 145 AA.
AC Q924Q6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHL186.2-D-J-C mu protein (Fragment).
GN Names=VHL186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR PIR; F28833; F28833.
DR PIR; F28833; F28833.
DR PIR; PH1105; PH1105.
DR PIR; PH1108; PH1108.
DR PIR; PH1114; PH1114.
DR PIR; PH1118; PH1118.
DR PIR; PH1119; PH1119.
DR PIR; PH1125; PH1125.
DR PIR; PH1126; PH1126.
DR PIR; PH1128; PH1128.
DR PIR; PH1129; PH1129.
DR PIR; PH1131; PH1131.
DR PIR; PH1134; PH1134.
DR PIR; PH1137; PH1137.
DR PIR; PH1139; PH1139.
DR PIR; PH1142; PH1142.
DR PIR; PH1144; PH1144.
DR PIR; PH1147; PH1147.
DR PIR; PH1149; PH1149.
DR PIR; PH1150; PH1150.
DR PIR; PH1151; PH1151.
DR PIR; PH1152; PH1152.
DR PIR; PH1153; PH1153.
DR HSSP; P01751; 1A6W.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
FT NON TER 1
FT NON TER 145
SQ SEQUENCE 145 AA; 16141 MW; 55A59A7908B2CD6A CRC64;

Query Match 47.8%; Score 44; DB 2; Length 145;
Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KIDPANGTKYDPIFOA 17
:||||:||||:|:|:
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 35
Q924Q9 PRELIMINARY; PRT; 145 AA.
AC Q924Q9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VHL186.2-D-J-C mu protein (Fragment).
GN Names=VHL186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;

```



```
FT NON_TER 1 1
FT NON_TER 145 145
SQ SEQUENCE 145 AA; 15996 MW; 35B1A36E4280BA81 CRC64;

Query Match      47.8%; Score 44; DB 2; Length 145;
Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17
   :|||:||||:|:
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 38
Q924R4
ID Q924R4 PRELIMINARY; PRT; 145 AA.
AC Q924R4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vh186.2-D-J-C mu protein (Fragment).
GN Name=Vh186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067797; BAB63270.1; -.
DR PIR; F28833; F28833.
DR PIR; F33932; F33932.
DR PIR; F33932; F33932.
DR PIR; PH1105; PH1105.
DR PIR; PH1108; PH1108.
DR PIR; PH1114; PH1114.
DR PIR; PH1118; PH1118.
DR PIR; PH1119; PH1119.
DR PIR; PH1125; PH1125.
DR PIR; PH1126; PH1126.
DR PIR; PH1128; PH1128.
DR PIR; PH1129; PH1129.
DR PIR; PH1131; PH1131.
DR PIR; PH1134; PH1134.
DR PIR; PH1137; PH1137.
DR PIR; PH1139; PH1139.
DR PIR; PH1142; PH1142.
DR PIR; PH1144; PH1144.
DR PIR; PH1147; PH1147.
DR PIR; PH1149; PH1149.
DR PIR; PH1150; PH1150.
DR PIR; PH1151; PH1151.
DR PIR; PH1152; PH1152.
DR PIR; PH1153; PH1153.
DR HSSP; P01751; IA6W.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 145 145
SQ SEQUENCE 145 AA; 16081 MW; ECDBIA135E05B8AA CRC64;

Query Match      47.8%; Score 44; DB 2; Length 145;
Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17
   :|||:||||:|:
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 39
Q924Q3
ID Q924Q3 PRELIMINARY; PRT; 146 AA.
AC Q924Q3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vh186.2-D-J-C mu protein (Fragment).
GN Name=Vh186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067797; BAB63282.1; -.
DR PIR; F28833; F28833.
DR PIR; F33932; F33932.
DR PIR; PH1105; PH1105.
DR PIR; PH1108; PH1108.
DR PIR; PH1114; PH1114.
DR PIR; PH1118; PH1118.
DR PIR; PH1119; PH1119.
DR PIR; PH1125; PH1125.
DR PIR; PH1126; PH1126.
DR PIR; PH1128; PH1128.
DR PIR; PH1129; PH1129.
DR PIR; PH1131; PH1131.
DR PIR; PH1134; PH1134.
DR PIR; PH1137; PH1137.
DR PIR; PH1139; PH1139.
DR PIR; PH1142; PH1142.
DR PIR; PH1144; PH1144.
DR PIR; PH1147; PH1147.
DR PIR; PH1149; PH1149.
DR PIR; PH1150; PH1150.
DR PIR; PH1151; PH1151.
DR PIR; PH1152; PH1152.
DR PIR; PH1153; PH1153.
DR HSSP; P01751; IA6W.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
FT NON_TER 1
FT NON_TER 146 146
SQ SEQUENCE 146 AA; 16136 MW; CEA8DD6E1955807F CRC64;

Query Match      47.8%; Score 44; DB 2; Length 146;
Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KIDPANGTKYDPIFOA 17
   :|||:||||:|:
Db 50 RIDPNSGGTKYNEKFKS 66

RESULT 40
Q924Q8
ID Q924Q8 PRELIMINARY; PRT; 146 AA.
AC Q924Q8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Vh186.2-D-J-C mu protein (Fragment).
GN Name=Vh186.2-D-J-C mu;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
```



This Page Blank (uspt)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 15:37:35 ; Search time 101.333 Seconds

(without alignments)  
41.984 Million cell updates/sec

Title: US-10-089-452-23

Perfect score: 65

Sequence: 1 PIYVASSWFAY 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	100.0	11	4	AAB86083 H. pylori
2	65	100.0	11	4	AAB86051 H. pylori
3	65	100.0	120	4	AAB86107 H. pylori
4	65	100.0	120	4	AAB86069 H. pylori
5	47	72.3	118	5	AAB29269 S. aureus
6	47	72.3	118	5	AAB29265 S. aureus
7	47	72.3	134	2	AAR12358 Heavy chain
8	47	72.3	134	2	AAR12326 Heavy chain
9	47	72.3	369	4	AAB73388 Anti-VHS
10	45	69.2	11	7	AD53469 MCP6 603
11	43	66.2	119	2	AAR01585 Lead bind
12	43	66.2	140	2	AAR12234 Mouse Mab
13	43	66.2	141	2	AAR12356 Heavy chain
14	41.5	63.8	137	6	ABG71659 HMPG-1 he
15	41	63.1	219	6	ABU44920 Protein e
16	41	63.1	232	6	ABU47897 Protein e
17	41	63.1	232	6	ABU47150 Protein e
18	41	63.1	253	2	AAR72599 Anti-dans
19	41	63.1	855	2	AAR43069 HIV-1 gp1
20	40	61.5	125	6	ABU15788 Protein e
21	40	61.5	366	6	ABM70196 Photorhab
22	40	61.5	402	3	AAY75127 Neisseria
23	40	61.5	402	3	AAY75128 Neisseria
24	40	61.5	2186	2	AAB31948 Mouse 22B
25	40	61.5	2858	4	ABG71150 Drosophila

26	40	61.5	3060	4	ABB58064	Abb58064 Drosophila
27	40	61.5	3672	2	AAB31950	Aab31950 Human bg
28	40	61.5	3788	2	AAW23594	Aaw23594 Murine Ly
29	40	61.5	3801	2	AAW31949	Aaw31949 Human bg
30	40	61.5	3801	8	ADR14706	Adr14706 Human NF-
31	39	60.0	13	8	ADL26969	Adl26969 Murine de
32	39	60.0	118	8	ADI26480	Adi26480 H. pylori
33	39	60.0	122	8	ADL26964	Adl26964 Mouse ant
34	39	60.0	298	4	AAU01247	Aau01247 B. subtil
35	39	60.0	405	2	AAW11151	Aaw11151 Chicken a
36	39	60.0	406	2	AAR60798	Aar60798 Chicken l
37	39	60.0	406	2	AAW15567	Aaw15567 Chicken l
38	39	60.0	451	8	ADN23872	Adn23872 Bacterial
39	39	60.0	520	8	ADS23804	Ads23804 Bacterial
40	39	60.0	744	3	AAB25567	Aab25567 Eucalyptu
41	39	60.0	1262	5	ABP74021	Abp74021 Candida a
42	38.5	59.2	115	7	ADJ80376	Adj80376 Hybrid hu
43	38.5	59.2	116	2	AAW04592	Aaw04592 Anti-DNA
44	38.5	59.2	116	7	ADJ80397	Adj80397 Hybrid hu
45	38	58.5	110	5	ABG71522	Abg71522 Mouse ant

## ALIGNMENTS

## RESULT 1

AAB86083

ID AAB86083 standard; peptide; 11 AA.

XX AC AAB86083;

DT 17-JUL-2001 (first entry)

DE H. pylori catalase derived antibody HP25/6m/1B5 heavy chain CDR3.

KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;  
KW acid-resistant microorganism; complementarity determining region; CDR;  
KW feces; heavy chain; light chain.

XX OS Unidentified.

XX WO200127612-A2.

XX PD 19-APR-2001.

XX PF 12-OCT-2000; 2000WO-EP010057.

XX PR 12-OCT-1999; 99EP-00120351.

XX PR 16-MAR-2000; 2000EP-00105592.

XX PR 31-MAR-2000; 2000EP-00107028.

XX PA 10-MAY-2000; 2000EP-00110110.

XX (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;

XX WPI; 2001-282086/29.

XX N-PSDB; AAF88110.

XX Detecting infections by acid-resistant microorganisms, particularly for  
diagnosing Helicobacter pylori, comprises immunochromatographic detection  
of antigen in feces.

XX Claim 26; Page 25; 90pp; German.

XX This invention describes a novel method for detecting infection by an  
acid-resistant microorganism (A), in a mammal, using  
immunochromatography. The method is used to diagnose infection by an acid  
-resistant microorganism (A), in a mammal, such as Helicobacter,  
Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,  
inexpensive and non-invasive, and may indicate the stage of infection. A  
test strip used in the method may include a filter to eliminate particles

CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence represents a  
 CC complementarity determining region (CDR) from an antibody raised against  
 CC the H. pylori catalase or beta-urease antigen which is used to illustrate  
 CC the method of the invention

XX Sequence 11 AA;

Query Match 100.0%; Score 65; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00014;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIYYASSWFAY 11  
 |||||  
 Db 1 PIYYASSWFAY 11

# RESULT 2

AAB86051

ID AAB86051 standard; peptide; 11 AA.

XX AAB86051;

DT 17-JUL-2001 (first entry)

DE H. pylori catalase derived antibody HP25/6m/1B5 heavy chain CDR3.

KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;

KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;

KW antibacterial; complementarity determining region.

XX Unidentified.

XX WO200127613-A2.

XX 19-APR-2001.

XX 12-OCT-2000; 2000WO-EP010058.

XX 12-OCT-1999; 99EP-00120351.

XX 16-MAR-2000; 2000EP-00105592.

XX 31-MAR-2000; 2000EP-00107028.

XX 10-MAY-2000; 2000EP-00110110.

XX (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX Reiter C, Cullmann G, Heppner P, Ringels A, Mueller H, Haindl E;

XX WPI; 2001-282087/29.

XX N-PSDB; AAF88053.

XX Detecting infections by acid-resistant microorganisms, particularly for

XX diagnosing Helicobacter pylori, comprises an immunoassay on a fecal

XX sample.

XX Claim 19; Page 16; 89pp; German.

XX This invention describes a novel method for detecting, in a mammal,

XX infection by an acid-resistant microorganism (A) which comprises reacting

XX a fecal sample with: (i) a receptor (R) such that a complex is formed

XX with an antigen (Ag) of (A); or (ii) two different R so that a three-part

XX complex is formed with Ag, and the formation of a complex detected. R are

CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an  
 CC infection. This sequence represents a complementarity determining region  
 CC (CDR) from an antibody generated against a Helicobacter pylori antigen  
 CC (catalase or beta-urease) which is used to illustrate the method of the  
 CC invention

XX Sequence 11 AA;

Query Match 100.0%; Score 65; DB 4; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00014;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 PIYYASSWFAY 11  
 |||||  
 Db 1 PIYYASSWFAY 11

# RESULT 3

AAB86107

ID AAB86107 standard; protein; 120 AA.

XX AC AAB86107;

DT 17-JUL-2001 (first entry)

DE H. pylori catalase antibody HP25/2m/1B5 V region heavy chain.

XX Catalase; beta-urease; antibody; antigen; detection; infection; epitope;

XX acid-resistant microorganism; complementarity determining region; CDR;

XX feces; heavy chain; light chain.

XX Unidentified.

XX WO200127612-A2.

XX 19-APR-2001.

XX 12-OCT-2000; 2000WO-EP010057.

XX 12-OCT-1999; 99EP-00120351.

XX 16-MAR-2000; 2000EP-00105592.

XX 31-MAR-2000; 2000EP-00107028.

XX 10-MAY-2000; 2000EP-00110110.

XX (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;

XX WPI; 2001-282086/29.

XX N-PSDB; AAF88153.

XX Detecting infections by acid-resistant microorganisms, particularly for

XX diagnosing Helicobacter pylori, comprises immunochromatographic detection

XX of antigen in feces.

XX Claim 31; Fig 3; 90pp; German.

XX This invention describes a novel method for detecting infection by an

XX acid-resistant microorganism (A), in a mammal, using

XX immunochromatography. The method is used to diagnose infection by an acid

XX resistant microorganism (A), in a mammal, such as Helicobacter,

XX Campylobacter or Mycobacterium, particularly H. pylori (most preferred),

XX H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,

XX inexpensive and non-invasive, and may indicate the stage of infection. A

XX test strip used in the method may include a filter to eliminate particles

XX present in the sample and only a single receptor provides a reasonably

XX secure diagnosis, with specificity and selectivity improved by detecting

XX several epitopes (of catalase) or different antigens (catalase and beta-

XX urease). The method can be automated. This sequence represents a



CC Helicobacter pylori catalase derived antibody V-region heavy chain  
 CC fragment used to illustrate the method of the invention  
 XX  
 SQ Sequence 120 AA;

Query Match 100.0%; Score 65; DB 4; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 0.0021;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PIYYASSWFAY 11  
 |||||  
 Db 99 PIYYASSWFAY 109

RESULT 4  
 AAB86069  
 ID AAB86069 standard; protein; 120 AA.  
 AC AAB86069;  
 XX  
 DT 17-JUL-2001 (first entry)  
 XX  
 DE H. pylori catalase antibody HP25/6m/1B5 V-region heavy chain.  
 XX Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
 KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
 KW antibacterial; complementarity determining region; V-region.  
 XX Unidentified.  
 OS  
 XX WO200127613-A2.  
 PN  
 XX 19-APR-2001.  
 PD  
 XX 12-OCT-2000; 2000WO-EP010058.  
 PF  
 XX 12-OCT-1999; 99EP-00120351.  
 PR 16-MAR-2000; 2000EP-0010592.  
 PR 31-MAR-2000; 2000EP-00107028.  
 PR 10-MAY-2000; 2000EP-00110110.  
 XX  
 PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
 XX  
 PI Reiter C, Cullmann G, Heppner P, Ringels A, Mueller H, Haendl E;  
 XX  
 DR WPI; 2001-282087/29.  
 DR N-PSDB; AAF88096.  
 XX  
 PT Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
 PT sample.  
 XX  
 PS Claim 27; Fig 3; 89pp; German.  
 XX  
 CC This invention describes a novel method for detecting, in a mammal,  
 CC infection by an acid-resistant microorganism (A) which comprises reacting  
 CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
 CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
 CC complex is formed with Ag, and the formation of a complex detected. R are  
 CC specific for an Ag which, after passage through the intestines, at least  
 CC in some mammals, retains a native (or corresponding) structure against  
 CC which the mammal produces antibodies (when immunized or infected with  
 CC (A), or its extracts, lysates or derived proteins (or fragments) or  
 CC synthetic peptides). The products of the invention have antibacterial  
 CC activity. The method is used to diagnose infection by Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the  
 CC progress of treatment. Receptors, particularly antibodies, directed  
 CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an

CC infection. This sequence represents a Helicobacter pylori anti-catalase  
 CC derived antibody HP25/6m/1B5 V-region heavy chain fragment which is  
 CC described in the method of the invention  
 XX  
 SQ Sequence 120 AA;

Query Match 100.0%; Score 65; DB 4; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 0.0021;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PIYYASSWFAY 11  
 |||||  
 Db 99 PIYYASSWFAY 109

RESULT 5  
 AAE29269  
 ID AAE29269 standard; protein; 118 AA.  
 XX  
 AC AAE29269;  
 XX  
 DT 27-JAN-2003 (first entry)  
 XX  
 DE S. aureus ClfA specific monoclonal antibody 35-220VHC-1 protein.  
 XX Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;  
 KW immunological; staphylococcal infection; impetigo; pneumonia; furuncle;  
 KW septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.  
 XX Staphylococcus aureus.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Region 31..35  
 FT /note= "CDR1"  
 FT Region 50..65  
 FT /note= "CDR2"  
 FT Region 98..107  
 FT /note= "CDR3"  
 XX  
 PN WO200272600-A2.  
 XX  
 PD 19-SEP-2002.  
 XX  
 PF 28-JAN-2002; 2002WO-US002296.  
 XX  
 PR 26-JAN-2001; 2001US-0264072P.  
 PR 12-MAR-2001; 2001US-0274611P.  
 PR 18-JUN-2001; 2001US-0298413P.  
 PR 30-JUL-2001; 2001US-0308116P.  
 XX  
 PA (INHI-) INHIBITEX INC.  
 XX  
 PI Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;  
 XX  
 DR WPI; 2002-759834/82.  
 DR N-PSDB; AAD46868.  
 XX  
 PT New anti-clumping factor A (ClfA) monoclonal antibody, useful for  
 PT treating or preventing Staphylococcus aureus infection e.g. wound  
 PT infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in  
 PT a human or animal.  
 XX  
 PS Claim 13; Page 37; 80pp; English.  
 XX  
 CC The invention relates to monoclonal antibody which binds the clumping  
 CC factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA  
 CC monoclonal antibody is useful for treating or preventing S. aureus  
 CC infection in a human or animal, and for inhibiting the binding of  
 CC staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment  
 CC S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3  
 CC protein is useful for inducing an immunological response in a human or  
 CC animal. These staphylococcal infections include wound infections, sepsis,  
 CC impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The

CC present sequence is Staphylococcus aureus ClfA specific monoclonal  
XX antibody 35-220VHC-1 (variable heavy sequence) protein  
SQ Sequence 118 AA;

Query Match 72.3%; Score 47; DB 5; Length 118;  
Best Local Similarity 77.8%; Pred. No. 2.7;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YYASSWFAY 11  
||:|||||  
Db 99 YYGNSWFAY 107

RESULT 6  
ID AAE29265 standard; protein; 118 AA.  
XX  
AC AAE29265;  
XX  
XX 27-JAN-2003 (first entry)  
XX  
XX S. aureus ClfA specific monoclonal antibody 13-2VHC-3 protein.  
XX  
XX Clumping factor A; ClfA; fibrinogen; fibrin; Clf40; Clf33; N3 protein;  
XX immunological; staphylococcal infection; impetigo; pneumonia; furuncle;  
XX septic arthritis; sepsis; endocarditis; antibacterial; wound; therapy.  
XX  
XX Staphylococcus aureus.  
XX  
XX  
XX Key Location/Qualifiers  
XX Region 31..35  
XX /note= "CDR1"  
XX Region 50..65  
XX /note= "CDR2"  
XX Region 98..107  
XX /note= "CDR3"  
XX  
XX WO200272600-A2.  
XX  
XX 19-SEP-2002.  
XX  
XX 28-JAN-2002; 2002WO-US002296.  
XX  
XX 26-JAN-2001; 2001US-0264072P.  
XX 12-MAR-2001; 2001US-0274611P.  
XX 18-JUN-2001; 2001US-0298413P.  
XX 30-JUL-2001; 2001US-0308116P.  
XX  
XX (NHI-) INHIBITEX INC.  
XX  
XX Patti JM, Hutchins JT, Domanski P, Patel P, Hall A;  
XX  
XX WPI; 2002-759834/82.  
XX N-PSDB; AAD46864.  
XX  
XX New anti-clumping factor A (ClfA) monoclonal antibody, useful for  
XX treating or preventing Staphylococcus aureus infection e.g. wound  
XX infections, impetigo, furuncles, pneumonia, septic arthritis or sepsis in  
XX a human or animal.  
XX  
XX Claim 13; Page 35; 80pp; English.  
XX  
XX The invention relates to monoclonal antibody which binds the clumping  
XX factor A (ClfA) protein from Staphylococcus aureus. The anti-ClfA  
XX monoclonal antibody is useful for treating or preventing S. aureus  
XX infection in a human or animal, and for inhibiting the binding of  
XX staphylococcal bacteria to fibrinogen or fibrin. The immunogenic fragment  
XX S. aureus Clf40 protein, S. aureus Clf33 protein, or the S. aureus N3  
XX protein is useful for inducing an immunological response in a human or  
XX animal. These staphylococcal infections include wound infections, sepsis,  
XX impetigo, furuncles, pneumonia, septic arthritis or endocarditis. The  
XX present sequence is Staphylococcus aureus ClfA specific monoclonal

CC antibody 13-2VHC-3 (variable heavy sequence) protein  
XX  
SQ Sequence 118 AA;

Query Match 72.3%; Score 47; DB 5; Length 118;  
Best Local Similarity 77.8%; Pred. No. 2.7;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YYASSWFAY 11  
||:|||||  
Db 99 YYGNSWFAY 107

RESULT 7  
ID AAR12358 standard; protein; 134 AA.  
XX  
XX AAR12358;  
XX  
XX 25-MAR-2003 (revised)  
XX 15-AUG-1991 (first entry)  
XX  
XX Heavy chain variable region of murine 1C11 immunoglobulin.  
XX  
XX Chimeric antibodies; immunoconjugates; HIV; AIDS.  
XX  
XX Mus musculus.  
XX  
XX WO9107493-A.  
XX  
XX 30-MAY-1991.  
XX  
XX 13-NOV-1989; 89US-00433730.  
XX  
XX 13-NOV-1989; 89US-00433730.  
XX  
XX (XOMA ) XOMA CORP.  
XX (GREC ) GREEN CROSS CORP.  
XX  
XX Better MD, Horwitz AH, Ghoshdaati P, Robinson R;  
XX  
XX WPI; 1991-178105/24.  
XX N-PSDB; AAQ12060.  
XX  
XX New chimeric mouse-human antibodies - used to detect, kill and remove HIV  
XX -1 antigen from sample.  
XX  
XX Disclosure; Fig 12; 107pp; English.  
XX  
XX This is the heavy-chain variable (V) region of a mouse mono-clonal  
XX antibody (MAB), 1C11, and is specific for an HIV-1 viral antigen. It is  
XX used in the construction of a chimeric MAB comprising heavy and light  
XX chains having murine V regions and human C regions. The chimeric MABs are  
XX more effective than murine MAB 1C11 since they have an increased  
XX compatibility in humans. The heavy and light chain V-regions are joined  
XX by manipulating their respective joining (J) regions, to generate  
XX immunconjugates, in association with e.g. toxins for HIV treatment. They  
XX can also be used in diagnosis of HIV. See also AAQ12056-59 and AAQ12061-  
XX 63. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003  
XX to correct PI field.) (Updated on 25-MAR-2003 to correct DR field.)  
XX  
XX Sequence 134 AA;

Query Match 72.3%; Score 47; DB 2; Length 134;  
Best Local Similarity 70.0%; Pred. No. 3.1;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IYYASSWFAY 11  
||:|||||  
Db 114 IYVTDWFAY 123

**RESULT 10**

ADFS3469  
ID ADFS3469 standard; peptide; 11 AA.  
XX  
AC ADFS3469;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE MCPC 603 antibody CDR3 mutant peptide amino acid sequence 12.  
XX  
KW walk-through mutagenesis; prototype amino acid; prototype nucleotide;  
XX mutant polypeptide production; MCPC 603;  
KW complementarity determining region; CDR; heavy chain; CDR3; mutant;  
KW mucin.  
XX  
OS Unidentified.  
OS Synthetic.  
OS  
XX WO2003089671-A1.  
PN  
XX 30-OCT-2003.  
PD  
XX 16-APR-2003; 2003WO-US011935.  
XX  
XX 17-APR-2002; 2002US-0373686P.  
PR  
XX (CREA/) CREA R.  
PA (CAPP/) CAPPUCCILLI G.  
PA  
XX Crea R, Cappuccilli G;  
PI  
XX WPI; 2003-854132/79.  
DR  
XX  
XX Walk-through mutagenesis of a nucleic acid encoding a polypeptide, useful  
PT for producing mutant polypeptides comprising synthesizing oligonucleotides  
PT comprising a nucleotide sequence for each target region of a prototype  
PT amino acid.  
XX  
XX Example; Fig 7; 40pp; English.  
PS  
XX This invention relates to a novel method of walk-through mutagenesis of a  
CC nucleic acid encoding a polypeptide which comprises synthesising a  
CC mixture of oligonucleotides comprising a nucleotide sequence for each  
CC target region of a prototype amino acid, where each oligonucleotide  
CC contains, at each sequence position in the target region, a prototype  
CC nucleotide for synthesis of the prototype amino acid, or a predetermined  
CC nucleotide that is required for synthesis of the predetermined amino  
CC acid. The method is useful in producing mutant polypeptides in which the  
CC overall presence of the predetermined amino acid is limited to one or two  
CC positions per mutated polypeptide, leaving the remaining amino acids in  
CC the targeted region intact or as close as possible to the prototype  
CC sequence. The invention was exemplified using the three complementarity  
CC determining regions (CDRs) of the heavy chain of the monoclonal antibody  
CC MCPC 603.  
XX  
XX  
SQ Sequence 11 AA;  
Query Match 69.2%; Score 45; DB 7; Length 11;  
Best Local Similarity 66.7%; Pred. No. 0.42;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 3 YYASSWFAY 11  
DB 2 YYSSWYSY 10  
||:||||:  
||:||||:  
RESULT 11  
AAW01585  
ID AAW01585 standard; protein; 119 AA.  
XX  
AC AAW01585;  
XX  
DT 22-AUG-1997 (first entry)  
XX

Lead binding MAB 11D11 heavy chain variable region.  
Monoclonal antibody; Fd fragment; lead cation; perfume; cosmetic;  
pharmaceutical; health care; skin treatment; pesticide; herbicide;  
heavy metal.  
Mus musculus.  
WO9639518-A1.  
12-DEC-1996.  
05-JUN-1996; 96WO-US009258.  
05-JUN-1995; 95US-00462798.  
10-OCT-1995; 95US-00541373.  
(BION-) BIONEERASKA INC.  
Wylie DE, Lopez O, Murray FJ;  
WPI; 1997-0431140/04.  
N-PSDB; AAT58259.  
DNA encoding heavy metal binding polypeptide sequences - used for  
detecting, removing, adding or neutralising heavy metals, such as lead  
cations.  
Claim 12; Page 73; 125pp; English.  
The present sequence represents the heavy chain variable region for  
monoclonal antibody (MAB) 11D11, which immunoreacts with a lead cation.  
The sequence was derived from RNA isolated from mouse hybridoma cells.  
The protein can be used for binding heavy metals, such as lead cations.  
It can be used for detecting, removing, adding or neutralising the heavy  
metals in biological and inanimate systems. It can be used in e.g.  
aqueous liquid systems, in biological or environmental systems or in such  
compositions as perfumes, cosmetics, pharmaceuticals, health care  
products, skin treatment products, pesticides, herbicides, solvents used  
in the production of semi-conductor and integrated circuit components and  
production materials for electronic components. The products can provide  
for applications involving minute amounts of specific heavy metals  
Sequence 119 AA;  
Query Match 66.2%; Score 43; DB 2; Length 119;  
Best Local Similarity 70.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 IYYASSWFAY 11  
DB 99 IYVGHLMFAY 108  
||| ||||  
||| ||||  
RESULT 12  
AAR12234  
ID AAR12234 standard; protein; 140 AA.  
XX  
AC AAR12234;  
XX  
DT 25-MAR-2003 (revised)  
DT 19-AUG-1991 (first entry)  
XX  
XX Mouse MAb 2G12 H chain V region.  
XX HIV-1; chimera.  
XX  
XX Mus sp.  
XX  
XX WO9107494-A.  
XX  
XX 30-MAY-1991.  
XX

PF 13-NOV-1989; 89US-00433703.  
 XX  
 PR 13-NOV-1989; 89US-00433703.  
 XX  
 PA (XOMA ) XOMA CORP.  
 PA (GREG ) GREEN CROSS CORP.  
 PA (ZOMA-) ZOMA CORP.  
 XX  
 PI Better MD, Horwitz AH, Ghoshdasti P, Robinson RR;  
 XX WPI; 1991-178106/24.  
 DR N-PSDB; AAQ12014.  
 DR  
 XX New chimeric mouse human antibodies - used in treatment, diagnosis and  
 PT prophylaxis of HIV infections.  
 PT  
 PS Disclosure; Fig 7; 108pp; English.  
 PS  
 CC The mouse VH gene product may be used to produce chimeric mouse- human  
 CC Abs against HIV-1 comprising human Ig constant regions and murine  
 CC variable regions. These novel sequence are useful in treatment, diagnosis  
 CC and prophylaxis of HIV infections, and may be produced by a bacterial,  
 CC yeast or mammalian expression system. (Updated on 25-MAR-2003 to correct  
 CC PA field.) (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 140 AA;  
 Query Match 66.2%; Score 43; DB 2; Length 140;  
 Best Local Similarity 60.0%; Pred. No. 16;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 IYYASSWFAY 11  
 :|||:||||  
 Db 121 LYGNFWFAY 130  
 RESULT 13  
 AAR12356  
 ID AAR12356 standard; protein; 141 AA.  
 XX  
 AC AAR12356;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 15-AUG-1991 (first entry)  
 DT  
 XX Heavy chain variable region of murine 2G12 immunoglobulin.  
 XX  
 XX Chimeric antibodies; immunoconjugates; HIV; AIDS.  
 XX  
 XX Mus musculus.  
 OS  
 XX WO9107493-A.  
 PN  
 XX 30-MAY-1991.  
 PD  
 XX 13-NOV-1989; 89US-00433730.  
 PF  
 XX 13-NOV-1989; 89US-00433730.  
 PR  
 XX (XOMA ) XOMA CORP.  
 PA (GREG ) GREEN CROSS CORP.  
 PA  
 XX Better MD, Horwitz AH, Ghoshdasti P, Robinson R;  
 PI WPI; 1991-178105/24.  
 XX DR N-PSDB; AAQ12058.  
 DR  
 XX New chimeric mouse-human antibodies - used to detect, kill and remove HIV  
 PT -1 antigen from sample.  
 PT  
 XX Disclosure; Fig 7; 107pp; English.  
 PS  
 XX This is the heavy-chain variable (V) region of a mouse monoclonal

CC antibody (MAB), 2G12, and is specific for an HIV-1 viral antigen. It is  
 CC used in the construction of a chimeric MAB comprising heavy and light  
 CC chains having murine V regions and human C regions. The chimeric MABs are  
 CC more effective than murine MAB 2G12 since they have an increased  
 CC compatibility in humans. The heavy and light chain V-regions are joined  
 CC by manipulating their respective joining (J) regions, to generate  
 CC restriction enzyme recognition sites. The chimeric MABs can be used as  
 CC immunoconjugates, in association with e.g. toxins for HIV treatment. They  
 CC can also be used in diagnosis of HIV. See also AAQ12056-57 and AAQ12059-  
 CC 63. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003  
 CC to correct PI field.) (Updated on 25-MAR-2003 to correct DR field.)  
 XX  
 SQ Sequence 141 AA;  
 Query Match 66.2%; Score 43; DB 2; Length 141;  
 Best Local Similarity 60.0%; Pred. No. 16;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 IYYASSWFAY 11  
 :|||:||||  
 Db 121 LYGNFWFAY 130  
 RESULT 14  
 ABG71659  
 ID ABG71659 standard; protein; 137 AA.  
 XX  
 AC ABG71659;  
 XX  
 DT 17-JAN-2003 (first entry)  
 DT  
 XX HMF-1 heavy chain variable region variant (Ala).  
 DE  
 XX Human; immunoglobulin variable domain; CDR1; framework region;  
 KW complementarity determining region 1; anti-MUC-1; HMF-1;  
 KW anti-mucin-1 monoclonal antibody; human milk fat globule-1; cancer;  
 KW tumour cell; mucin-like glycoprotein; neoplasm; carcinoma; breast cancer;  
 KW ovarian cancer; gastrointestinal cancer; epithelial-derived tumour;  
 KW heavy chain variable region; VH; cytostatic; mutant; mutein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 41  
 FT /note= "Substitution of wild-type Cys to Ala"  
 FT Region 50..54  
 FT /label= CDR1  
 FT Region 69..85  
 FT /label= CDR2  
 FT Misc-difference 115  
 FT /note= "Substitution of wild-type Cys to Ala"  
 FT Region 118..126  
 FT /label= CDR3  
 FT  
 XX WO200278613-A2.  
 PN  
 XX 10-OCT-2002.  
 PD  
 XX 02-APR-2002; 2002WO-US010304.  
 XX  
 PF 02-APR-2001; 2001US-0281182P.  
 PR  
 XX (PURD ) PURDUE PHARMA LP.  
 PA  
 XX Soltis DA, Burch RM, Shukla R;  
 PI WPI; 2003-040616/03.  
 XX DR N-PSDB; ABSS6088.  
 DR  
 XX New variant of an immunoglobulin variable domain comprising a  
 PT complementarity determining region (CDR) and framework regions flanking  
 PT the CDR, useful for treating or preventing e.g. breast cancer.

```
XX Example 1; Fig 3; 75pp; English.
XX
XX The present invention relates to a variant of an immunoglobulin variable
XX domain comprising at least one complementarity determining region (CDR)
XX and framework regions flanking the CDR, where the variant also comprises
XX a CDR region to which at least one binding sequence is added or
XX substituted, and flanking framework regions, where the binding sequence
XX is heterologous to the CDR and is an antigenic sequence from a mucin-1
XX (MUC-1) binding sequence. The variant immunoglobulin variable domain is
XX preferably a variant of the variable region of an anti-mucin monoclonal
XX antibody such as the anti-MUC-1 monoclonal antibody HMFG-1 (human milk
XX fat globule-1). Compositions comprising the variant immunoglobulin and
XX the methods of the present invention are useful for treating cancer by
XX eliciting an anti-idiotype response targeted to tumour cells that bear
XX mucin-like glycoproteins on neoplasms of epithelial origin, including
XX carcinomas of the breast, ovary and gastrointestinal tract. A vaccine
XX based on the sequence of the HMFG-1 monoclonal antibody may be used in
XX the treatment and prevention of breast cancer and other epithelial-
XX derived tumours. The present sequence represents HMFG-1 heavy chain
XX variable region variant (Ala)
XX
XX Sequence 137 AA;
XX
XX Query Match 63.8%; Score 41.5; DB 6; Length 137;
XX Best Local Similarity 53.3%; Pred. No. 28;
XX Matches 8; Conservative 2; Mismatches 0; Indels 5; Gaps 1;
XX
XX QY 2 IYVAS-----SWFAY 11
XX :|||||:|||||
XX Db 112 VYASRSYDFAWFAY 126
XX
XX RESULT 15
XX ABU44920
XX ID ABU44920 standard; protein; 219 AA.
XX
XX AC ABU44920;
XX
XX DT 19-JUN-2003 (first entry)
XX
XX DE Protein encoded by Prokaryotic essential gene #30447.
XX
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX OS Salmonella paratyphi.
XX
XX PN W0200277183-A2.
XX
XX PD 03-OCT-2002.
XX
XX PF 21-MAR-2002; 2002WO-US009107.
XX
XX PR 21-MAR-2001; 2001US-00815242.
XX
XX PR 06-SEP-2001; 2001US-00948993.
XX
XX PR 25-OCT-2001; 2001US-0342923P.
XX
XX PR 08-FEB-2002; 2002US-00072851.
XX
XX PR 06-MAR-2002; 2002US-0362699P.
XX
XX PA (ELIT-) ELITRA PHARM INC.
XX
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI: 2003-029926/02.
XX
XX DR N-PSDB; ACA48790.
XX
XX XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 72844; 1766pp; English.
XX
```

```
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 219 AA;
XX
XX Query Match 63.1%; Score 41; DB 6; Length 219;
XX Best Local Similarity 55.6%; Pred. No. 59;
XX Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 PIYVASSWF 9
XX :|||:|:
XX Db 69 PLYGSEWY 77
XX
XX RESULT 16
XX ABU47897
XX ID ABU47897 standard; protein; 232 AA.
XX
XX AC ABU47897;
XX
XX DT 19-JUN-2003 (first entry)
XX
XX DE Protein encoded by Prokaryotic essential gene #33424.
XX
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX OS Salmomella typhi.
XX
XX PN W0200277183-A2.
XX
XX PD 03-OCT-2002.
XX
XX PF 21-MAR-2002; 2002WO-US009107.
XX
XX PR 21-MAR-2001; 2001US-00815242.
XX
XX PR 06-SEP-2001; 2001US-00948993.
XX
XX PR 25-OCT-2001; 2001US-0342923P.
XX
XX PR 08-FEB-2002; 2002US-00072851.
XX
XX PR 06-MAR-2002; 2002US-0362699P.
XX
XX PA (ELIT-) ELITRA PHARM INC.
XX
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
```



XX Single chain antibody; dansyl; SCAB; bispecific; tetravalent;  
 KW immunoglobulin; heavy chain; light chain; primer.

XX Synthetic.  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH 126..140  
 FT Peptide /note="Flexible linker"

XX W05909917-A1.

XX 13-APR-1995.

XX 07-OCT-1994; 94WO-US011411.

XX 07-OCT-1993; 93US-00134556.

XX (REGC ) UNIV CALIFORNIA.

XX Morrison SL, Coloma MJ;

XX WPI; 1995-155258/20.

XX N-PSDB; AAQ86755.

XX Prodn. of bi-specific, tetra-valent antibodies - by recombinant  
 PT expression of DNA encoding variable heavy and light immunoglobulin  
 FT chains.

XX Example 1; Page 19; 33pp; English.

XX The anti-dansyl single chain antibody was produced using PCR to join two  
 CC variable regions by overlap extension. A flexible linker region  
 CC consisting of three gly-gly-gly-ser repeats was used to connect the C  
 CC -terminus of the heavy chain variable region (Vh) and the N-terminus of  
 CC the light chain variable region (Vl) in order to maximise flexibility  
 CC while minimising steric hindrance during protein folding. For the PCR  
 CC reaction, the 3' primer for Vh (AAQ86751) and the 5' primer for Vl  
 CC (AAQ86752) were designed to have a complementary region in the linker  
 CC portion. The Vh and Vl regions were independently amplified, followed by  
 CC a second PCR reaction in which these products were used as templates. The  
 CC complementary sequences in the linker were hybridised, joining the V  
 CC regions and two primers (AAQ86753, AAQ86754) were then used to amplify  
 CC the single chain antibody. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 253 AA;

Query Match 63.1%; Score 41; DB 2; Length 253;

Best Local Similarity 70.0%; Pred. No. 69;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 IYYASSWFAY 11

Db 105 IYHYFPWFAY 114

RESULT 19

AAW43069

ID AAW43069 standard; peptide; 855 AA.

XX AAW43069;

XX 17-OCT-2003 (revised)

DT 11-SEP-1998 (first entry)

XX HIV-1 gp120 protein fragment from isolate SC.

XX gp120 protein; purification; fractionation; ion exchange; chromatography;  
 KW binding affinity; CD4; hydrophobic interaction; size exclusion; vaccine.  
 XX Human immunodeficiency virus 1.

PN US5696238-A.

XX 09-DEC-1997.

XX 11-MAY-1995; 95US-00439286.

XX 20-AUG-1991; 91US-00684963.

PR 16-AUG-1993; 93US-00109002.

PR 09-MAY-1994; 94US-00240073.

XX (CHIR ) CHIRON CORP.

XX Scandella C, Haigwood NL;

XX WPI; 1998-041353/04.

XX Purification of HIV gp120 - using chromatographic methods.

XX Disclosure; Fig 2A-W; 53pp; English.

XX AAW43066-W43080 are fragments of the gp120 protein from different human  
 CC immunodeficiency virus type I (HIV-1) isolates. These proteins are used  
 CC in a novel method for purifying HIV gp120 so as to provide a purified  
 CC gp120 glycopeptide having protein/protein binding properties  
 CC substantially identical to natural viral HIV gp120. The method involves  
 CC fractionating a crude gp120 preparation containing full-length,  
 CC glycosylated gp120 using ion exchange chromatography so as to provide a  
 CC first collection of fractions. A fraction from the first collection is  
 CC selected that exhibits specific binding affinity for CD4 peptide, thereby  
 CC producing a first fractionated material. The first fractionated material  
 CC is fractionated by hydrophobic interaction chromatography so as to  
 CC provide a second collection of fractions from which a second collection  
 CC is selected that exhibits specific binding affinity for CD4 peptide. This  
 CC second fraction is fractionated by size exclusion chromatography so as to  
 CC provide a third collection of fractions exhibiting specific binding  
 CC affinity for CD4 peptide, thereby providing the purified gp120. The  
 CC purified gp120 can be used for antibody production and in vaccines.  
 CC (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 855 AA;

Query Match 63.1%; Score 41; DB 2; Length 855;

Best Local Similarity 60.0%; Pred. No. 2.7e+02;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PIYVASSWFA 10

Db 217 PIHYCARWFA 226

RESULT 20

ABU15788

ID ABU15788 standard; protein; 125 AA.

XX ABU15788;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #1315.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Salmonella typhimurium.

XX W0200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.





```

XX PN WO9957280-A2.
XX PD 11-NOV-1999.
XX PF 30-APR-1999; 99WO-US009346.
XX PR 01-MAY-1998; 98US-0083758P.
XX PR 31-JUL-1998; 98US-0094869P.
XX PR 02-SEP-1998; 98US-0098994P.
XX PR 02-SEP-1998; 98US-0099062P.
XX PR 09-OCT-1998; 98US-0103749P.
XX PR 09-OCT-1998; 98US-0103794P.
XX PR 09-OCT-1998; 98US-0103796P.
XX PR 25-FEB-1999; 99US-0121528P.
XX PA (CHIR ) CHIRON CORP.
XX PA (GENO-) INST GENOMIC RES.
XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
XX PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX PI Tettelin H, Venter JC;
XX XX
XX DR WPI: 2000-062150/05.
XX DR N-PSDB; AAZ53889.
XX XX
XX PT Novel Neisserial polypeptides predicted to be useful antigens for
XX PT vaccines and diagnostics.
XX PS Claim 2; Page 885; 1453pp; English.
XX CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
XX CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
XX CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
XX CC PCR primers used in the exemplification of the present invention. The
XX CC polypeptides, the polynucleotides, antibodies and compositions of the
XX CC invention can be used as vaccines, as diagnostic reagents, and as
XX CC immunogenic compositions. The polypeptides can be used in the manufacture
XX CC of medicaments for treating or preventing infection due to Neisserial
XX CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
XX CC Neisseria bacteria, or to raise antibodies. They may also be used to
XX CC screen for agonists or antagonists, which may themselves have use as
XX CC antibacterial agents. The polynucleotides of the invention may also be
XX CC used in gene therapy protocols
XX SQ Sequence 402 AA;

Query Match 61.5%; Score 40; DB 3; Length 402;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYYASSWFA 10
Db 29 PVFFAMGWFA 38
|:::| |||
|:::| |||

RESULT 23
AAZ75128
ID AAZ75128 standard; protein; 402 AA.
XX AC AAZ75128;
XX XX
XX DT 21-MAR-2000 (first entry)
XX DE
XX DE Neisseria meningitidis ORF 574 protein sequence SEQ ID NO:1730.
XX XX
XX KW Neisseria meningitidis; Neisseria gonorrhoea; antigen; vaccine;
XX KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
XX KW antibacterial; gene therapy.
XX OS
XX OS Neisseria meningitidis.
XX XX
XX PN WO9957280-A2.

```

```

XX PD 11-NOV-1999.
XX PF 30-APR-1999; 99WO-US009346.
XX PR 01-MAY-1998; 98US-0083758P.
XX PR 31-JUL-1998; 98US-0094869P.
XX PR 02-SEP-1998; 98US-0098994P.
XX PR 02-SEP-1998; 98US-0099062P.
XX PR 09-OCT-1998; 98US-0103749P.
XX PR 09-OCT-1998; 98US-0103794P.
XX PR 09-OCT-1998; 98US-0103796P.
XX PR 25-FEB-1999; 99US-0121528P.
XX PA (CHIR ) CHIRON CORP.
XX PA (GENO-) INST GENOMIC RES.
XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
XX PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX PI Tettelin H, Venter JC;
XX XX
XX DR WPI: 2000-062150/05.
XX DR N-PSDB; AAZ53890.
XX XX
XX PT Novel Neisserial polypeptides predicted to be useful antigens for
XX PT vaccines and diagnostics.
XX PS Claim 2; Page 886; 1453pp; English.
XX CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
XX CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
XX CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
XX CC PCR primers used in the exemplification of the present invention. The
XX CC polypeptides, the polynucleotides, antibodies and compositions of the
XX CC invention can be used as vaccines, as diagnostic reagents, and as
XX CC immunogenic compositions. The polypeptides can be used in the manufacture
XX CC of medicaments for treating or preventing infection due to Neisserial
XX CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
XX CC Neisseria bacteria, or to raise antibodies. They may also be used to
XX CC screen for agonists or antagonists, which may themselves have use as
XX CC antibacterial agents. The polynucleotides of the invention may also be
XX CC used in gene therapy protocols
XX SQ Sequence 402 AA;

Query Match 61.5%; Score 40; DB 3; Length 402;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYYASSWFA 10
Db 29 PVFFAMGWFA 38
|:::| |||
|:::| |||

RESULT 24
AAW31948
ID AAW31948 standard; protein; 2186 AA.
XX AC AAW31948;
XX XX
XX DT 14-APR-1998 (first entry)
XX DE
XX DE Mouse 22B/30B (candidate beige) protein.
XX XX
XX KW Chediak-Higashi syndrome; beige; bg gene; 22B/30B gene;
XX KW intracellular vesicle; mouse; therapy; diagnosis.
XX OS
XX OS Mus musculus.
XX PN WO9734914-A1.
XX PD 25-SEP-1997.
XX XX

```

```

PF 21-MAR-1997; 97WO-US005068.
XX
FR 22-MAR-1996; 96US-0013883P.
FR 19-APR-1996; 96US-0015673P.
PR 01-JUL-1996; 96US-0021064P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA (UTAH ) UNIV UTAH.
XX
PI Kaplan J, Perou CM, Moore KJ;
XX
XX WPI; 1997-502728/46.
DR N-PSDB; AAT89255.
XX
XX Murine and human long and short form beige genes - useful to screen for
PT compounds to treat, or to diagnose intracellular vesicle disorders,
PT specifically Chediak-Higashi syndrome.
XX
XX Claim 9; Fig 4; 133pp; English.
XX
XX This polypeptide is encoded by the novel murine 22B/30B gene (see
CC AAT89255), which represents a candidate beige (bg) gene involved in the
CC normal differentiation and/or function of intracellular vesicles. The
CC amino acid sequence of the predicted bg gene product indicates that the
CC protein is novel. The human bg gene (see AAT89256 and AAT89257) has also
CC been identified. These genes, when mutated, are responsible for human
CC Chediak-Higashi syndrome (CHS). An intracellular vesicle disorder,
CC specifically CHS, can be diagnosed by measuring bg gene expression in a
CC patient sample, e.g. by detecting mRNA transcripts of the bg gene, the bg
CC gene product or a bg gene mutation contained in the genome of the mammal,
CC specifically in a splice site of the bg gene (claimed). Genetically
CC engineered host cells can be used to screen for compounds useful for the
CC treatment of intracellular vesicle disorders, specifically CHS, e.g. a
CC molecule that modulates the expression of the bg gene in a mammal
CC (claimed)
XX
XX Sequence 2186 AA;
SQ
Query Match 61.5%; Score 40; DB 2; Length 2186;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PIYVASSW 8
Db 1325 PIYPTSW 1332

RESULT 25
ABB71150
ID ABB71150 standard; protein; 2858 AA.
XX
XX ABB71150;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 40242.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
KW
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR N-PSDB; ABL02167.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 40242; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2858 AA;
SQ
Query Match 61.5%; Score 40; DB 4; Length 2858;
Best Local Similarity 45.5%; Pred. No. 1.6e+03;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PIYVASSW 11
Db 236 PMFFANSWWNY 246

RESULT 26
ABB58064
ID ABB58064 standard; protein; 3060 AA.
XX
XX ABB58064;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 984.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
KW
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR N-PSDB; ABL02167.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 984; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and

```

CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
 CC ABBS72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 3060 AA;  
 Query Match 61.5%; Score 40; DB 4; Length 3060;  
 Best Local Similarity 45.5%; Pred. No. 1.7e+03;  
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 PIYVASSWFAY 11  
 |:::|::|  
 Db 236 PMPFANSWMNY 246  
 RESULT 27  
 AAW31950  
 ID AAW31950 standard; protein; 3672 AA.  
 XX  
 AC AAW31950;  
 XX  
 DT 14-APR-1998 (first entry)  
 XX  
 DE Human bg protein associated with Chediak-Higashi syndrome.  
 XX  
 KW Chediak-Higashi syndrome; beige; bg gene; intracellular vesicle; human;  
 KW therapy; diagnosis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Location/Qualifiers  
 FT Peptide 3585..3579  
 /note= "WD40 or G protein-beta subunit repeat motif"  
 XX  
 WO9734914-A1.  
 XX  
 PD 25-SEP-1997.  
 XX  
 PF 21-MAR-1997; 97WO-US005068.  
 XX  
 PR 22-MAR-1996; 96US-0013883P.  
 PR 19-APR-1996; 96US-0015673P.  
 PR 01-JUL-1996; 96US-0021064P.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 PA (UTAH ) UNIV UTAH.  
 XX  
 PI Kaplan J, Perou CM, Moore KJ;  
 XX  
 DR WPI; 1997-502728/46.  
 DR N-PSDB; AAT89257.  
 XX  
 PT Murine and human long and short form beige genes - useful to screen for  
 PT compounds to treat, or to diagnose intracellular vesicle disorders,  
 PT specifically Chediak-Higashi syndrome.  
 XX  
 PS Claim 9; Fig 8; 133pp; English.  
 XX  
 CC This polypeptide is encoded by an alternatively spliced transcript of the  
 CC human beige (bg) gene (see AAT89257). The amino acid sequence of the  
 CC predicted bg gene product indicates that the protein is novel. A longer,  
 CC putatively full-length bg transcript (see AAT89256) encodes a 3801-amino  
 CC acid protein (see AAW31949). The bg gene product is involved in the  
 CC normal differentiation and/or function of intracellular vesicles.  
 CC Mutation of the human bg gene is responsible for human Chediak-Higashi  
 CC syndrome (CHS). An intracellular vesicle disorder, specifically CHS, can  
 CC be diagnosed by measuring bg gene expression in a patient sample, e.g. by  
 CC detecting mRNA transcripts of the bg gene, the bg gene product or a bg  
 CC gene mutation contained in the genome of the mammal, specifically in a

CC splice site of the bg gene (claimed). Genetically engineered host cells  
 CC can be used to screen for compounds useful for the treatment of  
 CC intracellular vesicle disorders, specifically CHS, e.g. a molecule that  
 CC modulates the expression of the bg gene in a mammal (claimed)  
 XX  
 SQ Sequence 3672 AA;  
 Query Match 61.5%; Score 40; DB 2; Length 3672;  
 Best Local Similarity 75.0%; Pred. No. 2.1e+03;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 PIYVASSW 8  
 |:::|::|  
 Db 2811 PIYPTSW 2818  
 RESULT 28  
 AAW23594  
 ID AAW23594 standard; protein; 3788 AA.  
 XX  
 AC AAW23594;  
 XX  
 DT 10-FEB-1998 (first entry)  
 XX  
 DE Murine Lyst1 long isoform.  
 XX  
 KW Lyst1; mouse; lysosomal trafficking regulator; Chediak-Higashi syndrome;  
 KW CH syndrome; autoimmune disease; tumour; vaccine; therapy; diagnosis.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO9728262-A1.  
 XX  
 PD 07-AUG-1997.  
 XX  
 PF 31-JAN-1997; 97WO-US001748.  
 XX  
 PR 01-FEB-1996; 96US-0011146P.  
 PR 20-DEC-1996; 96US-0033599P.  
 PR 23-DEC-1996; 96US-0034346P.  
 XX  
 PA (UYFL ) UNIV FLORIDA.  
 XX  
 PI Kingsmore SF, Barbosa-Alleyne MDFS;  
 XX  
 DR WPI; 1997-402616/37.  
 DR N-PSDB; AAT74199.  
 XX  
 PT Mammalian lysosomal trafficking regulators LYST1, Lyst1, LYST2 and Lyst2  
 PT - useful to diagnose Chediak-Higashi syndrome.  
 XX  
 PS Claim 3; Page 101-103; 237pp; English.  
 XX  
 CC This protein comprises the longer isoform of murine lysosomal trafficking  
 CC regulator Lyst1, a protein that regulates degradation of lysosomes,  
 CC late endosomes and acidic secretory granules primarily in leukocytes. Its  
 CC sequence was deduced from a Lyst1 cDNA clone (AAT74199) that had been  
 CC identified from a beige (bg) mutation critical region yeast artificial  
 CC chromosome. Alternative splicing also yields a shorter isoform (see  
 CC AAW23595) of Lyst1. Homologous human sequences (see AAW23596 and  
 CC AAW23597) have been identified that are associated with Chediak-Higashi  
 CC syndrome (CHS). Murine Lyst1 and human LYST1 polypeptides (see AAW23593-  
 CC 96) can be expressed in recombinant host cells for use in raising  
 CC antibodies, in vaccines and in the development of therapeutic approaches  
 CC to treatment of autoimmune diseases and certain types of tumours. The  
 CC existence of an animal model of CHS with a similar genetic lesion will  
 CC assist efforts to develop novel therapies for this disease  
 XX  
 SQ Sequence 3788 AA;  
 Query Match 61.5%; Score 40; DB 2; Length 3788;  
 Best Local Similarity 75.0%; Pred. No. 2.2e+03;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 PIYYASSW 8
Db 2927 PIYYPTSW 2934

RESULT 29
AAW31949
ID AAW31949 standard; protein; 3801 AA.
XX AC AAW31949;
XX DT 14-APR-1998 (first entry)
XX DE Human bg protein associated with Chediak-Higashi syndrome.
XX KW Chediak-Higashi syndrome; beige; bg gene; intracellular vesicle; human;
XX KW therapy; diagnosis.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 3694..4708
XX FT /note= "WD40 or G protein-beta subunit repeat motif"
XX FT Misc-difference 3767
XX FT /note= "encoded by TNC"
XX FN W09734914-A1.
XX PN 25-SEP-1997.
XX PD 21-MAR-1997; 97WO-US005068.
XX PP 22-MAR-1996; 96US-0013883P.
XX PR 19-APR-1996; 96US-0015673P.
XX PR 01-JUL-1996; 96US-0021064P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PA (UTAH) UNIV UTAH.
XX PI Kaplan J, Perou CM, Moore KJ;
XX PI WPI; 1997-502728/46.
XX DR N-PSDB; AAT89256.
XX DR Murine and human long and short form beige genes - useful to screen for
XX FT compounds to treat, or to diagnose intracellular vesicle disorders,
XX FT specifically Chediak-Higashi syndrome.
XX PS Claim 9; Fig 7; 133pp; English.
XX CC This polypeptide is encoded by the novel human beige (bg) gene (see
XX CC AAT89256) that is involved in the normal differentiation and/or function
XX CC of intracellular vesicles. The amino acid sequence of the predicted bg
XX CC gene product indicates that the protein is novel. An alternatively
XX CC spliced transcript of the human bg gene (see AAT89257) encodes a 3672-
XX CC amino acid protein (see AAW31950). Mutation of the human bg gene is
XX CC responsible for human Chediak-Higashi syndrome (CHS). An intracellular
XX CC vesicle disorder, specifically CHS, can be diagnosed by measuring bg gene
XX CC expression in a patient sample, e.g. by detecting mRNA transcripts of the
XX CC bg gene, the bg gene product or a bg gene mutation contained in the
XX CC genome of the mammal, specifically in a splice site of the bg gene
XX CC (claimed). Genetically engineered host cells can be used to screen for
XX CC compounds useful for the treatment of intracellular vesicle disorders,
XX CC specifically CHS, e.g. a molecule that modulates the expression of the bg
XX CC gene in a mammal (claimed)
XX SQ Sequence 3801 AA;

Query Match 61.5%; Score 40; DB 2; Length 3801;
Best Local Similarity 75.0%; Pred. No. 2.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASSW 8
Db 2940 PIYYPTSW 2947

RESULT 30
ADRI4706
ID ADRI4706 standard; protein; 3801 AA.
XX AC ADRI4706;
XX DT 21-OCT-2004 (first entry)
XX DE Human NF-kappaB pathway-associated protein SeqID778.
XX KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
XX KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;
XX KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
XX KW immunosuppressive; vulnery; gene therapy; immune disorder;
XX KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
XX KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
XX KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
XX KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
XX KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
XX KW viral replication; host cell survival; evasion of immune response;
XX KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
XX KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAS;
XX KW autoimmune disorder; hyper immune activity;
XX KW aberrant acute phase response; hypercongenital condition; birth defect;
XX KW necrotic lesion; wound; organ transplant rejection;
XX KW aberrant signal transduction; proliferating disorder; cancer;
XX KW HIV propagation; human.
XX OS Homo sapiens.
XX PN W02004065577-A2.
XX PD 05-AUG-2004.
XX PF 13-JAN-2004; 2004WO-US000798.
XX PR 14-JAN-2003; 2003US-0440068P.
XX PR 12-MAY-2003; 2003US-0469757P.
XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX PI Nadler SG, Neubauer MG, Feder JN, Carman J;
XX PI WPI; 2004-562168/54.
XX DR N-PSDB; ADRI4705.
XX CC New isolated polynucleotides and polypeptides associated with NF-kappaB
XX CC pathway, useful for diagnosing, treating, or preventing disorders or
XX CC diseases associated with NF-kappaB pathway.
XX PS Claim 6; SEQ ID NO 778; 237pp; English.
XX CC This invention relates to the novel association of protein sequences (and
XX CC the genes which encode them) to the NF-kappaB pathway. The invention may
XX CC be useful for the production of compounds with an antiinflammatory,
XX CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
XX CC gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,
XX CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
XX CC vulnery activity or for gene therapy. The proteins and nucleotides are
XX CC useful for diagnosing, preventing, treating, or ameliorating conditions
XX CC or diseases associated with the NF-kappaB pathway. The condition is an
XX CC immune disorder, an inflammatory disorder, an inflammatory disorder
XX CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
XX CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
XX CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
XX CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
XX CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
XX CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory

```

CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick  
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper  
 CC immune activity, disorders related to aberrant acute phase responses,  
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,  
 CC organ transplant rejection, conditions related to organ transplant  
 CC rejection, disorders related to aberrant signal transduction,  
 CC proliferating disorders, cancers and HIV propagation in cells infected  
 CC with other viruses. The present sequence is that of a human protein which  
 CC is subject to the novel association with the NF-kappaB pathway of the  
 CC invention. Note: This sequence does not appear in the specification but  
 CC was obtained by the indexer from Genbank.

XX Sequence 3801 AA;

Query Match 61.5%; Score 40; DB 8; Length 3801;

Best Local Similarity 75.0%; Pred. No. 2.2e+03;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYVASSW 8

Db 2940 PIYVPTSW 2947  
 ||||| :||

RESULT 31

ADL26969 ID ADL26969 standard; peptide; 13 AA.

XX AC ADL26969;

XX DT 20-MAY-2004 (first entry)

XX DE Murine derived human CDR3 transplant antibody peptide.

XX KW human chimeric antibody; complementarity determining region; CDR;

XX KW transplant antibody; prostate gland specific membrane antigen; PSMA;

XX KW cytosolic; cancer; prostate; mouse; murine.

XX OS Mus musculus.

XX PN JP2004000045-A.

XX PD 08-JAN-2004.

XX PF 31-MAY-2002; 2002JJP-00158727.

XX PR 31-MAY-2002; 2002JJP-00158727.

XX PA (KYOW ) KYOWA HAKKO KOGYO KK.

XX DR WPI; 2004-286376/27.

XX Novel human chimeric or complementarity determining region transplant  
 PT antibody specific for prostate gland specific membrane antigen, useful  
 PT as diagnostic and therapeutic agent of cancer.

PS Claim 4; SEQ ID NO 9; 62pp; Japanese.

XX The invention relates to a novel human chimeric antibody or human  
 CC complementarity determining region (CDR) transplant antibody, specific  
 CC for prostate gland specific membrane antigen (PSMA), each comprising of  
 CC a heavy chain consisting of variable heavy chain (VH) and constant heavy  
 CC chain (CH) and a light chain (L) consisting of variable light chain (VL)  
 CC and constant light chain (CL) or its fragment. The novel antibodies have  
 CC cytosolic activity. The human chimeric antibody or transplant antibody  
 CC is useful as a diagnostic agent or a therapeutic agent for cancer. The  
 CC human chimeric antibody or transplant antibody is effective in an assay  
 CC of PSMA. The human chimeric antibody or transplant antibody can  
 CC effectively diagnose or treat cancer, specifically prostate cancer. This  
 CC sequence represents a human complementarity determining region (CDR)  
 CC transplant antibody peptide, derived from a mouse, of the invention.

XX Sequence 13 AA;

XX SQ Sequence 118 AA;

Query Match 60.0%; Score 39; DB 8; Length 118;

Best Local Similarity 66.7%; Pred. No. 65;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Query Match

60.0%; Score 39; DB 8; Length 13;

Best Local Similarity 72.7%; Pred. No. 5.5;

Matches 8; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy 3 YVASS--WFAY 11

Db 3 YGSSHWWFAY 13

|||||

RESULT 32

AD126480 ID AD126480 standard; protein; 118 AA.

XX AC AD126480;

XX DT 15-APR-2004 (first entry)

XX DE H. pylori HpU-2-H SEQ ID NO:16.

XX KW antibody; enzyme; virucide; anti-HIV; cytostatic; antibacterial;

XX KW helicobacter pylori urease inhibitor;

XX KW chemokine receptor CCR-5 antagonist; cancer; infectious disease;

XX KW Helicobacter pylori; HIV.

XX OS Helicobacter pylori.

XX PN WO2004009805-A1.

XX PD 29-JAN-2004.

XX PF 18-JUL-2003; 2003WO-JP009147.

XX PR 19-JUL-2002; 2002JP-00211756.

XX PR 19-JUL-2002; 2002JP-00211768.

XX PR 27-FEB-2003; 2003JP-00051943.

XX PR 17-JUL-2003; 2003JP-00198270.

XX PR 17-JUL-2003; 2003JP-00198281.

XX PR 17-JUL-2003; 2003JP-00198292.

XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX PI Uda T, Hifumi E;

XX DR WPI; 2004-132963/13.

XX DR N-PSDB; ADI26513.

XX Screening potential antibody enzymes by identification of a catalytic  
 PT triptet residue in the stereostructure for production of antibody enzymes  
 PT as diagnostic and therapeutic agents for cancer and infectious diseases  
 PT including HIV infection.

XX Claim 24; SEQ ID NO 16; 232pp; Japanese.

XX The invention relates to a novel method for producing antibody enzymes  
 CC comprising a structural analysis step which confirms the existence in the  
 CC predicted stereostructure of the antibody based on its amino acid  
 CC sequence of a catalytic triptet residue structure in which a serine  
 CC residue, an aspartic acid residue, and a histidine or glutamic acid  
 CC residue are located sterically close to one another. An enzyme produced  
 CC by the method of the invention has virucide, anti-HIV, cytostatic, and  
 CC antibacterial activity. The enzyme acts as a helicobacter pylori urease  
 CC inhibitor, or chemokine receptor CCR-5 (HIV target) antagonist. The  
 CC method of the invention is useful for the treatment, prevention and  
 CC diagnosis of cancer and infectious diseases, especially infection by  
 CC Helicobacter pylori or HIV. The present sequence is used in the  
 CC exemplification of the invention.







ADN23872  
ID ADN23872 standard; protein; 451 AA.  
XX  
AC ADN23872;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polypeptide #6525.  
XX  
KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
OS Bacteria.  
XX  
PN US2003233675-A1.  
XX  
PD 18-DEC-2003.  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
PR 21-FEB-2002; 2002US-0360039P.  
XX  
PA (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX  
DR WPI; 2004-061375/06.  
XX  
PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
PS Claim 1; SEQ ID NO 6525; 122pp; English.  
XX  
CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 451 AA;  
Query Match 60.0%; Score 39; DB 8; Length 451;  
Best Local Similarity 63.6%; Pred. No. 2.9e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIYASSWFAY 11  
|||: |||  
Db 180 PIMYSCSWPAY 190  
RESULT 39  
ADS22804  
ID ADS22804 standard; protein; 520 AA.  
XX  
AC ADS22804;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polypeptide #11837.  
XX  
KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
OS Bacteria.  
XX  
PN US2003233675-A1.  
XX  
PD 18-DEC-2003.  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
PR 21-FEB-2002; 2002US-0360039P.  
XX  
PA (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX  
DR WPI; 2004-061375/06.  
XX  
PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
PS Claim 1; SEQ ID NO 11837; 122pp; English.  
XX  
CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX

CC format from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 520 AA;

Query Match 60.0%; Score 39; DB 8; Length 520;  
Best Local Similarity 63.6%; Pred. No. 3.5e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYASSWFAY 11  
| | | | |  
Db 126 PIFASDWEAY 136

RESULT 40

AAB25567  
ID AAB25567 standard; protein; 744 AA.

XX AC AAB25567;

XX DT 27-NOV-2000 (first entry)

XX DE Eucalyptus grandis cell signalling involved protein SEQ ID NO:942.

XX KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
plant cell signalling; modulation; transgenic plant; pathogen; growth;  
environmental change; development; cell proliferation; differentiation;  
elongation; survival; disease resistance; nutrient metabolism.

XX OS Eucalyptus grandis.

XX PN MO200042171-A1.

XX PD 20-JUL-2000.

XX PF 11-JAN-2000; 2000WO-US000724.

XX PR 12-JAN-1999; 99US-00228986.

XX PR 01-NOV-1999; 99US-0162866P.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI Strabala TJ, Nieuwenhuizen NJ;

XX WPI; 2000-476052/41.

Isolated polynucleotide encoding a polypeptide involved in cell signaling  
used for generating transgenic plants with modified responses to external  
signals.

PS Claim 3; Page 520-522; 527pp; English.

XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide  
and protein sequences isolated from eucalyptus (Eucalyptus grandis) or  
pine (Pinus radiata also known as Monterey pine). The protein sequences  
are involved in cell signalling. The polynucleotide and protein sequences  
can be used to modify the response of plant cells to external signals  
e.g. environmental changes or pathogens during the growth and development  
of a plant. They can be used to modify cell proliferation,  
differentiation, elongation and survival, resistance to disease and  
nutrient metabolism. Examples of modifications which can be produced are  
altered fruit ripening and senescence of leaves and flowers e.g. to delay  
senescence and prolong the life of cut flowers or enhance senescence of  
reproductive organs to engineer sterile plants. Other modifications can  
be used to delay senescence in selected cell types or organs providing  
fruit and vegetables which have a longer shelf life between harvest and  
consumption, or to decrease branching frequency in forest tree species  
giving long stretches of valuable knot-free clear wood which can be used  
in solid timber furniture and veneers

XX Sequence 744 AA;

Query Match 60.0%; Score 39; DB 3; Length 744;  
Best Local Similarity 75.0%; Pred. No. 5.2e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 YASSWFAY 11  
| | | | |  
Db 242 YINSWFAY 249

Search completed: June 3, 2005, 15:56:46  
Job time : 103.333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 15:49:21 ; Search time 25.6667 Seconds  
(without alignments)  
31.992 Million cell updates/sec

Title: US-10-089-452-23

Perfect score: 65

Sequence: 1 PIYVASSWFAY 11

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	66.2	119	3	US-08-767-128-20
2	41	63.1	855	3	Sequence 20, Appl
3	40	61.5	143	4	Sequence 14, Appl
4	40	61.5	143	4	Sequence 31985, A
5	40	61.5	311	4	Sequence 47202, A
6	40	61.5	2186	2	Sequence 16680, A
7	40	61.5	2186	3	Sequence 2, Appl
8	40	61.5	3672	2	Sequence 12, Appl
9	40	61.5	3672	3	Sequence 12, Appl
10	40	61.5	3801	2	Sequence 10, Appl
11	40	61.5	3801	3	Sequence 10, Appl
12	39	60.0	405	1	Sequence 2, Appl
13	39	60.0	406	3	Sequence 2, Appl
14	39	60.0	744	4	Sequence 942, App
15	38.5	59.2	116	3	Sequence 19, Appl
16	38	58.5	409	4	Sequence 24, Appl
17	38	58.5	449	1	Sequence 1, Appl
18	38	58.5	449	3	Sequence 1, Appl
19	38	58.5	449	4	Sequence 1, Appl
20	38	58.5	524	4	Sequence 20256, A
21	37.5	57.7	118	2	Sequence 74, Appl
22	37.5	57.7	118	3	Sequence 2, Appl
23	37.5	57.7	118	3	Sequence 2, Appl
24	37	56.9	10	2	Sequence 25, Appl
25	37	56.9	10	4	Sequence 25, Appl
26	37	56.9	13	3	Sequence 7, Appl
27	37	56.9	61	4	Sequence 5956, Ap

28 37 56.9 119 2 US-08-318-157B-2 Sequence 2, Appli  
29 37 56.9 119 2 US-08-318-157B-8 Sequence 8, Appli  
30 37 56.9 119 2 US-08-318-157B-9 Sequence 9, Appli  
31 37 56.9 119 2 US-08-318-157B-10 Sequence 10, Appli  
32 37 56.9 119 2 US-08-318-157B-11 Sequence 11, Appli  
33 37 56.9 119 2 US-08-318-157B-12 Sequence 12, Appli  
34 37 56.9 119 2 US-08-318-157B-13 Sequence 13, Appli  
35 37 56.9 119 2 US-08-318-157B-14 Sequence 14, Appli  
36 37 56.9 119 2 US-08-318-157B-15 Sequence 15, Appli  
37 37 56.9 119 2 US-08-318-157B-17 Sequence 17, Appli  
38 37 56.9 119 2 US-08-318-157B-57 Sequence 57, Appli  
39 37 56.9 119 2 US-08-318-157B-58 Sequence 58, Appli  
40 37 56.9 119 4 US-09-253-794-2 Sequence 2, Appli  
41 37 56.9 119 4 US-09-253-794-8 Sequence 8, Appli  
42 37 56.9 119 4 US-09-253-794-9 Sequence 9, Appli  
43 37 56.9 119 4 US-09-253-794-10 Sequence 10, Appli  
44 37 56.9 119 4 US-09-253-794-11 Sequence 11, Appli  
45 37 56.9 119 4 US-09-253-794-12 Sequence 12, Appli

#### ALIGNMENTS

RESULT 1  
US-08-767-128-20  
; Sequence 20, Application US/08767128  
; Patent No. 6111079  
; GENERAL INFORMATION:  
; APPLICANT: WYLIE, DWANE E.  
; APPLICANT: LOPEZ, OSVALDO  
; APPLICANT: MURRAY, PETER JOSEPH  
; APPLICANT: GOEBEL, PETER  
; TITLE OF INVENTION: LEAD BINDING POLYPEPTIDES AND  
; TITLE OF INVENTION: NUCLEOTIDES CODING THEREFORE  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
; STREET: 3100 No. 6111079west Center, 90 South Seventh St  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/767,128  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE: 04-DEC-1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/09258  
; FILING DATE: 05-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/541,373  
; FILING DATE: 10-OCT-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/462,798  
; FILING DATE: 05-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carter, Charles G.  
; REGISTRATION NUMBER: 35,093  
; REFERENCE/DOCKET NUMBER: 8648.49USF1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612/371-5278  
; TELEFAX: 612/332-9081  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 20:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-767-128-20

Query Match 66.2%; Score 43; DB 3; Length 119;
Best Local Similarity 70.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 IYVASSWFAY 11
   |||
Db 99 IYGHLMFAY 108
   |||

RESULT 2
US-07-956-483-14
; Sequence 14, Application US/07956483
; Patent No. 6261799
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; TITLE OF INVENTION: GP160 VARIANT
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/956,483
; FILING DATE: 31-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/19742
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 05392
; FILING DATE: 02-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 017753-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 855 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-956-483-14

Query Match 63.1%; Score 41; DB 3; Length 855;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PIYVASSWFA 10
   |||

; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEtical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-10-089-452-23.ra1

Query Match 61.5%; Score 40; DB 4; Length 143;
Best Local Similarity 45.5%; Pred. No. 59;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 PIYVASSWFA 11
   |||
Db 33 PMFFANSWNY 43
   |||

RESULT 4
US-09-270-767-47202
; Sequence 47202, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47202
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47202

Query Match 61.5%; Score 40; DB 4; Length 143;
Best Local Similarity 45.5%; Pred. No. 59;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 PIYVASSWFA 11
   |||
Db 33 PMFFANSWNY 43
   |||

RESULT 5
US-09-248-796A-16680
; Sequence 16680, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
```

; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 16680  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-16680

Query Match 61.5%; Score 40; DB 4; Length 311;  
Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIYVASSWF 9  
|:|:|:|:|  
Db 271 PMPFASNWF 279

RESULT 6  
US-08-822-445-2  
; Sequence 2, Application US/08822445  
; Patent No. 5952223  
; GENERAL INFORMATION:  
; APPLICANT: Kaplan, Jerry  
; APPLICANT: Perou, Charles  
; APPLICANT: Moore, Karen  
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS  
; TITLE OF INVENTION: AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/822.445  
; FILING DATE: 21-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-062-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2186 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-822-445-2

Query Match 61.5%; Score 40; DB 2; Length 2186;  
Best Local Similarity 75.0%; Pred. No. 8.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYVASSW 8  
|:|:|:|:|  
Db 1325 PIYPTSW 1332

RESULT 7  
US-09-396-540-2  
; Sequence 2, Application US/09396540  
; Patent No. 6310182

; GENERAL INFORMATION:  
; APPLICANT: Kaplan, Jerry  
; APPLICANT: Perou, Charles  
; APPLICANT: Moore, Karen  
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS  
; TITLE OF INVENTION: AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/396,540  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/822,445  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-062-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2186 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-09-396-540-2

Query Match 61.5%; Score 40; DB 3; Length 2186;  
Best Local Similarity 75.0%; Pred. No. 8.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYVASSW 8  
|:|:|:|:|  
Db 1325 PIYPTSW 1332

RESULT 8  
US-08-822-445-12  
; Sequence 12, Application US/08822445  
; Patent No. 5952223  
; GENERAL INFORMATION:  
; APPLICANT: Kaplan, Jerry  
; APPLICANT: Perou, Charles  
; APPLICANT: Moore, Karen  
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS  
; TITLE OF INVENTION: AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible

```
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,445
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3672 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-822-445-12

Query Match 61.5%; Score 40; DB 2; Length 3672;
Best Local Similarity 75.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASSW 8
Db 2811 PIYYPTSW 2818

RESULT 9
US-09-396-540-12
; Sequence 12, Application US/09396540
; Patent No. 6310182
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/396,540
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/822,445
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3672 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-396-540-12

Query Match 61.5%; Score 40; DB 3; Length 3672;
Best Local Similarity 75.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASSW 8
Db 2811 PIYYPTSW 2818

RESULT 10
US-08-822-445-10
; Sequence 10, Application US/08822445
; Patent No. 5952223
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,445
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3801 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-822-445-10

Query Match 61.5%; Score 40; DB 2; Length 3801;
Best Local Similarity 75.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASSW 8
Db 2940 PIYYPTSW 2947

RESULT 11
US-09-396-540-10
; Sequence 10, Application US/09396540
; Patent No. 6310182
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Perou, Charles
; APPLICANT: Moore, Karen
```

;; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS  
;; TITLE OF INVENTION: AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME  
;; NUMBER OF SEQUENCES: 32  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Pennie & Edmonds  
;; STREET: 1155 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036/2711  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/396,540  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/822,445  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Coruzzi, Laura A.  
;; REGISTRATION NUMBER: 30,742  
;; REFERENCE/DOCKET NUMBER: 7953-062-999  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-790-9090  
;; TELEFAX: 212-869-8864  
;; TELEX: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 10:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 3801 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
;; US-09-396-540-10

Query Match 61.5%; Score 40; DB 3; Length 3801;  
Best Local Similarity 75.0%; Pred. No. 1.4e+03;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASW 8  
Db 2940 PIYYTSW 2947

RESULT 12  
US-08-406-070-2  
; Sequence 2, Application US/08406070  
; Patent No. 5610063  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Daniel S.  
; APPLICANT: Walker, John C.  
; TITLE OF INVENTION: A cDNA For  
; TITLE OF INVENTION: alpha-N-acetyl-galactosaminidase From Gallus domesticus  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Reising, Ethington, Barnard, Perry & Milton  
; STREET: P.O. Box 4390  
; CITY: Troy  
; STATE: Michigan  
; COUNTRY: US  
; ZIP: 48099  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/406,070  
; FILING DATE:  
; CLASSIFICATION: 435

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kohn, Kenneth I.  
;; REGISTRATION NUMBER: 30,955  
;; REFERENCE/DOCKET NUMBER: UMO P-317  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (810) 689-3500  
;; TELEFAX: (810) 689-4071  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 405 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-406-070-2

Query Match 60.0%; Score 39; DB 1; Length 405;  
Best Local Similarity 63.6%; Pred. No. 2.3e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYYASW 11  
Db 166 PIVYCSW 176

RESULT 13  
US-09-070-356-2  
; Sequence 2, Application US/09070356  
; Patent No. 6228631  
; GENERAL INFORMATION:  
; APPLICANT: Alex Zhu  
; APPLICANT: Jack Goldstein  
; TITLE OF INVENTION: Recombinant a-N-  
; TITLE OF INVENTION: Acetylglactosaminidase  
; TITLE OF INVENTION: Enzyme and cDNA Encoding  
; TITLE OF INVENTION: Said Enzyme  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amster, Rothstein & Ebenstein  
; STREET: 90 Park Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Word Processor (ASCII)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/070,356  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/037,248  
; FILING DATE: March 26, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pasqualini, Patricia A.  
; REGISTRATION NUMBER: 34,894  
; REFERENCE/DOCKET NUMBER: 63475/12  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 697-5995  
; TELEFAX: (212) 286-0854 or 286-0082  
; TELEX: TWX 710-581-4766  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 406  
; TYPE: amino acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA to mRNA  
; DESCRIPTION: CDNA to mRNA  
; HYPOTHETICAL: no

ANTI-SENSE: yes  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM: chicken liver  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE: library  
POSITION IN GENOME: unknown  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY: chicken liver a-N-  
NAME/KEY: acetylglucosaminidase  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-09-070-356-2

Query Match 60.0%; Score 39; DB 3; Length 406;  
Best Local Similarity 63.6%; Pred. No. 2.3e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYVASSWPAY 11  
Db 167 PIYVCSWPAY 177

RESULT 14  
US-10-101-464A-942  
; Sequence 942, Application US/10101464A  
; Patent No. 6768041  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Higgins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; FILE REFERENCE: 11000.1020c2  
; CURRENT FILING DATE: 2002-03-18  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR FILING DATE: 09/228,986  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR FILING DATE: 60/162,866  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR FILING DATE: PCT/US00/00724  
; NUMBER OF SEQ ID NOS: 989  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 942  
; LENGTH: 744  
; TYPE: PRT  
; ORGANISM: Eucalyptus grandis  
US-10-101-464A-942

Query Match 60.0%; Score 39; DB 4; Length 744;  
Best Local Similarity 75.0%; Pred. No. 4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 YASSWPAY 11  
Db 242 YINSWPAY 249

RESULT 15  
US-08-881-037-19  
; Sequence 19, Application US/08881037  
; Patent No. 6080588  
; GENERAL INFORMATION:  
; APPLICANT: Glick, Gary D.  
; APPLICANT: Swanson, Patrick C.  
; TITLE OF INVENTION: DNA BINDING ANTIBODIES  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/881,037  
; FILING DATE: 23-JUN-1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/443,540  
; FILING DATE: 18-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kanski, Antoinette F.  
; REGISTRATION NUMBER: 34,202  
; REFERENCE/DOCKET NUMBER: 203442110710  
; TELEPHONE: (650) 813-5600  
; TELEFAX: (650) 494-0792  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 116 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-881-037-19

Query Match 59.2%; Score 38.5; DB 3; Length 116;  
Best Local Similarity 58.3%; Pred. No. 81;  
Matches 7; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 PIYVAS-SWPAY 11  
Db 94 PAYISNYPWPAY 105

RESULT 16  
US-09-560-761B-24  
; Sequence 24, Application US/09560761B  
; Patent No. 6787683  
; GENERAL INFORMATION:  
; APPLICANT: Dellapenna, Dean  
; APPLICANT: Collakova, Eva  
; APPLICANT: Coughlan, Sean J.  
; APPLICANT: Heintjaris, Timothy G.



;; TITLE OF INVENTION: PHYTYL/PRENYLTRANSFERASE NUCLEIC ACIDS,  
;; FILE REFERENCE: 1095R  
;; CURRENT APPLICATION NUMBER: US/09/560,761B  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: 09/307,460  
;; PRIOR FILING DATE: 1999-05-07  
;; NUMBER OF SEQ ID NOS: 32  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 24  
;; LENGTH: 409  
;; TYPE: PRT  
;; ORGANISM: Glycine max  
US-09-560-761B-24

Query Match 58.5%; Score 38; DB 4; Length 409;  
Best Local Similarity 66.7%; Pred. No. 3.2e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 IYVASSWFA 10  
|||:||||:  
Db 32 IYFASSWIS 40

RESULT 17  
US-08-570-157-1  
;; Sequence 1, Application US/08570157  
;; Patent No. 5750353  
;; GENERAL INFORMATION:  
;; APPLICANT: Kopin, Alan S.  
;; APPLICANT: Beinborn, Martin  
;; TITLE OF INVENTION: ASSAY FOR NON-PEPTIDE AGONISTS TO  
;; TITLE OF INVENTION: PEPTIDE HORMONE RECEPTORS  
;; NUMBER OF SEQUENCES: 23  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Fish & Richardson P.C.  
;; STREET: 225 Franklin Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02110-2804  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/570,157  
;; FILING DATE: 11-DEC-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Clark, Paul T.  
;; REGISTRATION NUMBER: 30,162  
;; REFERENCE/DOCKET NUMBER: 00398/109001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617/542-5070  
;; TELEFAX: 617/542-8906  
;; TELEX: 200154  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 449 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: not relevant  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-570-157-1

Query Match 58.5%; Score 38; DB 1; Length 449;  
Best Local Similarity 54.5%; Pred. No. 3.5e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYVASSWFAY 11  
|||:|:|:

Db 350 PIYSANTWCAF 360  
RESULT 18  
US-09-076-510-1  
;; Sequence 1, Application US/09076510  
;; Patent No. 6376198  
;; GENERAL INFORMATION:  
;; APPLICANT: Kopin, Alan S.  
;; APPLICANT: Beinborn, Martin  
;; TITLE OF INVENTION: ASSAY FOR NON-PEPTIDE AGONISTS TO  
;; TITLE OF INVENTION: PEPTIDE HORMONE RECEPTORS  
;; NUMBER OF SEQUENCES: 23  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Clark & Eibing LLP  
;; STREET: 176 Federal Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02110  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: FastSeq version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/076,510  
;; FILING DATE: 12 MAY 1998  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/570,157  
;; FILING DATE: 11-DEC-1995  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Clark, Paul T.  
;; REGISTRATION NUMBER: 30,162  
;; REFERENCE/DOCKET NUMBER: 00398/109002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617/428-0200  
;; TELEFAX: 617/438-7045  
;; TELEX:

;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 449 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: not relevant  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-09-076-510-1

Query Match 58.5%; Score 38; DB 3; Length 449;  
Best Local Similarity 54.5%; Pred. No. 3.5e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYVASSWFAY 11  
|||:|:|:  
Db 350 PIYSANTWCAF 360

RESULT 19  
US-09-004-349-1  
;; Sequence 1, Application US/09004349A  
;; Patent No. 6566080  
;; GENERAL INFORMATION:  
;; APPLICANT: Kopin, Alan S.  
;; APPLICANT: Beinborn, Martin  
;; TITLE OF INVENTION: ASSAY FOR AND USES OF PEPTIDE HORMONE  
;; TITLE OF INVENTION: RECEPTOR AGONISTS  
;; FILE REFERENCE: 00398/118002  
;; CURRENT APPLICATION NUMBER: US/09/004,349A  
;; CURRENT FILING DATE: 1998-01-08  
;; EARLIER APPLICATION NUMBER: 08/570,157  
;; EARLIER FILING DATE: 1995-12-11  
;; NUMBER OF SEQ ID NOS: 23



```
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-767-128-2

Query Match          57.7%; Score 37.5; DB 3; Length 118;
Best Local Similarity 46.7%; Pred. NO. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

Qy 2 IYYA-----SWFAY 11
Db 93 VYICARRGNSAWFAY 107

RESULT 23
US-07-987-264-14
; Sequence 14, Application US/07987264
; Patent No. 6204366
; GENERAL INFORMATION:
; APPLICANT: VERHOEVEN, MARTINE ELISA
; TITLE OF INVENTION: SPECIFIC BINDING AGENTS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/987,264
; FILING DATE: 08-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9019553.8
; FILING DATE: 07-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB PCT/GB91/01511
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 200232/P3095USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 118 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-987-264-14

Query Match          57.7%; Score 37.5; DB 3; Length 118;
Best Local Similarity 46.7%; Pred. NO. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

Qy 2 IYYAS-----SWFAY 11
Db 93 VYICRSYDFAWFAY 107

RESULT 24
US-08-318-157B-25
; Sequence 25, Application US/08318157B

; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-318-157B-25

Query Match          56.9%; Score 37; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IYYASSWFAY 11
Db 1 LYFGFPWFAY 10

RESULT 25
US-09-253-794-25
; Sequence 25, Application US/09253794
; Patent No. 6676924
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/253,794
```

Query Match 56.9%; Score 37; DB 3; Length 13;

APPLICANT: HANSEN, Hans J.  
APPLICANT: ARMOUR, Kathryn L.  
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318,157B  
FILING DATE: 05-OCT-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: SAXE, Bernhard D.  
REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 18733/464  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-318-157B-2

Query Match 56.9%; Score 37; DB 2; Length 119;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IYYASSWFAY 11  
Db 99 LYFGFPWFAY 108

RESULT 29  
US-08-318-157B-8  
Sequence 8, Application US/08318157B  
Patent No. 5874540  
GENERAL INFORMATION:  
APPLICANT: HANSEN, Hans J.  
APPLICANT: ARMOUR, Kathryn L.  
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318,157B  
FILING DATE: 05-OCT-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:

NAME: SAXE, Bernhard D.  
REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 18733/464  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-318-157B-8

Query Match 56.9%; Score 37; DB 2; Length 119;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IYYASSWFAY 11  
Db 99 LYFGFPWFAY 108

RESULT 30  
US-08-318-157B-9  
Sequence 9, Application US/08318157B  
Patent No. 5874540  
GENERAL INFORMATION:  
APPLICANT: HANSEN, Hans J.  
APPLICANT: ARMOUR, Kathryn L.  
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318,157B  
FILING DATE: 05-OCT-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: SAXE, Bernhard D.  
REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 18733/464  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-318-157B-9

Query Match 56.9%; Score 37; DB 2; Length 119;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IYYASSWFAY 11

```
Db          :|: ||||
           99 LYFGFPWFAY 108

RESULT 31
US-08-318-157B-10
; Sequence 10, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; APPLICANT: ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-318-157B-11

Query Match      56.9%; Score 37; DB 2; Length 119;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      2 IYVASSWFAY 11
       :|: ||||
Db      99 LYFGFPWFAY 108

RESULT 33
US-08-318-157B-12
; Sequence 12, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; APPLICANT: ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-318-157B-10

Query Match      56.9%; Score 37; DB 2; Length 119;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      2 IYVASSWFAY 11
       :|: ||||
Db      99 LYFGFPWFAY 108

RESULT 32
US-08-318-157B-11
; Sequence 11, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; APPLICANT: ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
```

```
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-318-157B-12

Query Match          56.9%; Score 37; DB 2; Length 119;
Best Local Similarity 50.0%; Pred. NO. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IYASSWFAY 11
Db 99 LYFGFPWFAY 108

RESULT 34
US-08-318-157B-13
; Sequence 13, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; APPLICANT: ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; MOLECULE TYPE: MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-318-157B-13

Query Match          56.9%; Score 37; DB 2; Length 119;
Best Local Similarity 50.0%; Pred. NO. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IYASSWFAY 11
Db 99 LYFGFPWFAY 108

RESULT 35
US-08-318-157B-14
; Sequence 14, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; APPLICANT: ARMOUR, Kathryn L.
```

```
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; MOLECULE TYPE: MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-318-157B-14

Query Match          56.9%; Score 37; DB 2; Length 119;
Best Local Similarity 50.0%; Pred. NO. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IYASSWFAY 11
Db 99 LYFGFPWFAY 108

RESULT 36
US-08-318-157B-15
; Sequence 15, Application US/08318157B
; Patent No. 5874540
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; APPLICANT: ARMOUR, Kathryn L.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
; MOLECULE TYPE: MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
```

REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 18733/464  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-318-157B-15

Query Match 56.9%; Score 37; DB 2; Length 119;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IYYASSWFAY 11  
:|: ||||  
Db 99 LYFGFPWFAY 108

RESULT 37  
US-08-318-157B-17  
; Sequence 17, Application US/08318157B  
; Patent No. 5874540  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; APPLICANT: ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/318,157B  
; FILING DATE: 05-OCT-1994  
; CLASSIFICATION: 424  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665  
; REFERENCE/DOCKET NUMBER: 18733/464  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 119 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-318-157B-17

Query Match 56.9%; Score 37; DB 2; Length 119;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IYYASSWFAY 11  
:|: ||||  
Db 99 LYFGFPWFAY 108

RESULT 38  
US-08-318-157B-57  
; Sequence 57, Application US/08318157B  
; Patent No. 5874540  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; APPLICANT: ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/318,157B  
; FILING DATE: 05-OCT-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SAXE, Bernhard D.  
; REGISTRATION NUMBER: 28,665  
; REFERENCE/DOCKET NUMBER: 18733/464  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 57:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 119 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-318-157B-57

Query Match 56.9%; Score 37; DB 2; Length 119;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IYYASSWFAY 11  
:|: ||||  
Db 99 LYFGFPWFAY 108

RESULT 39  
US-08-318-157B-58  
; Sequence 58, Application US/08318157B  
; Patent No. 5874540  
; GENERAL INFORMATION:  
; APPLICANT: HANSEN, Hans J.  
; APPLICANT: ARMOUR, Kathryn L.  
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED  
; TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible



```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-318-157B-58

Query Match          56.9%; Score 37; DB 2; Length 119;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      2 IYYASSWFAY 11
       :|: |||
Db      99 LYFGFPWFAY 108

Search completed: June 3, 2005, 16:03:56
Job time : 26.6667 secs

; MOLECULE TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-253-794-2

Query Match          56.9%; Score 37; DB 4; Length 119;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      2 IYYASSWFAY 11
       :|: |||
Db      99 LYFGFPWFAY 108

Search completed: June 3, 2005, 16:03:56
Job time : 26.6667 secs

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,157B
; FILING DATE: 05-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-318-157B-58

Query Match          56.9%; Score 37; DB 2; Length 119;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      2 IYYASSWFAY 11
       :|: |||
Db      99 LYFGFPWFAY 108

RESULT 40
US-09-253-794-2
; Sequence 2, Application US/09253794
; Patent No. 6676924
; GENERAL INFORMATION:
; APPLICANT: HANSEN, Hans J.
; TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CEA HUMANIZED
;                   MOUSE MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/253,794
; FILING DATE: 22-Feb-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,157
; FILING DATE: 05-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 18733/464
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 16:01:46 ; Search time 79.6667 Seconds  
(without alignments)  
47.730 Million cell updates/sec

Title: US-10-089-452-23

Perfect score: 65

Sequence: 1 PIYVASSWFAY 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1465611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 1465611

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	73.8	200	15	US-10-425-114-55157
2	47	72.3	118	14	US-10-056-052-8
3	47	72.3	118	14	US-10-056-052-16
4	45	69.2	11	15	US-10-417-895A-57
5	42	64.6	11	15	US-10-417-895A-59
6	41.5	63.8	118	15	US-10-435-614-11
7	41.5	63.8	118	15	US-10-435-614-13
8	41.5	63.8	137	16	US-10-469-125-39
9	41.5	63.8	177	15	US-10-435-614-21
10	41.5	63.8	260	15	US-10-435-614-20
11	41	63.1	219	15	US-10-282-122A-72844
12	41	63.1	232	15	US-10-282-122A-75074
13	41	63.1	232	15	US-10-282-122A-75821

14	40	61.5	125	15	US-10-282-122A-43712	Sequence 43712, A
15	40	61.5	206	15	US-10-425-114-47000	Sequence 47000, A
16	40	61.5	376	14	US-10-156-761-14201	Sequence 14201, A
17	40	61.5	2174	14	US-10-087-887-87	Sequence 87, Appl
18	40	61.5	2186	9	US-09-927-668-2	Sequence 2, Appl
19	40	61.5	3672	9	US-09-927-668-12	Sequence 12, Appl
20	40	61.5	3801	9	US-09-927-668-10	Sequence 10, Appl
21	40	61.5	3801	16	US-10-755-889-777	Sequence 777, Appl
22	39.5	60.8	14	10	US-09-791-551-97	Sequence 97, Appl
23	39.5	60.8	146	10	US-09-791-551-111	Sequence 111, Appl
24	39	60.0	81	15	US-10-424-599-254210	Sequence 254210,
25	39	60.0	118	16	US-10-437-963-140199	Sequence 140199,
26	39	60.0	216	17	US-10-887-231-21	Sequence 21, Appl
27	39	60.0	298	17	US-10-984-449-30	Sequence 30, Appl
28	39	60.0	451	15	US-10-369-493-6525	Sequence 6525, Ap
29	39	60.0	520	15	US-10-369-493-11837	Sequence 11837, A
30	39	60.0	727	16	US-10-437-963-185152	Sequence 185152,
31	39	60.0	744	14	US-10-101-464A-942	Sequence 942, App
32	39	60.0	744	17	US-10-864-252-942	Sequence 942, App
33	39	60.0	1262	14	US-10-032-585-7858	Sequence 7858, Ap
34	38.5	59.2	115	15	US-10-308-817-136	Sequence 136, App
35	38.5	59.2	115	15	US-10-453-698-136	Sequence 136, App
36	38.5	59.2	116	15	US-10-308-817-157	Sequence 157, App
37	38.5	59.2	116	15	US-10-453-698-157	Sequence 157, App
38	38	58.5	34	16	US-10-437-963-130365	Sequence 130365,
39	38	58.5	74	15	US-10-424-599-283720	Sequence 283720,
40	38	58.5	76	15	US-10-424-599-234514	Sequence 234514,
41	38	58.5	110	16	US-10-471-475A-4	Sequence 4, Appli
42	38	58.5	164	16	US-10-471-475A-23	Sequence 23, Appl
43	38	58.5	376	16	US-10-767-701-45444	Sequence 45444, A
44	38	58.5	449	14	US-10-127-940-1	Sequence 1, Appli
45	38	58.5	449	14	US-10-441-757-1	Sequence 1, Appli

#### ALIGNMENTS

##### RESULT 1

US-10-425-114-55157  
; Sequence 55157, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 55157  
; LENGTH: 200  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73043H11\_FLI.pep  
US-10-425-114-55157

Query Match 73.8%; Score 48; DB 15; Length 200;  
Best Local Similarity 72.7%; Pred. No. 9.5;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYVASSWFAY 11

Db 17 PIYVSSWFY 27

##### RESULT 2

US-10-056-052-8

; Sequence 8, Application US/10056052  
; Publication No. US20030099656A1  
; GENERAL INFORMATION:  
; APPLICANT: PATTI, Joseph M  
; APPLICANT: HUTCHINS, Jeff T  
; APPLICANT: DOMANSKI, Paul  
; APPLICANT: PATEL, Pratiksha  
; APPLICANT: HALL, Andrea  
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN . . .  
; FILE REFERENCE: P07069US04/BAS  
; CURRENT APPLICATION NUMBER: US/10/056,052  
; CURRENT FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/308,116  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/298,413  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/274,611  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 60/264,072  
; PRIOR FILING DATE: 2001-01-26  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 118  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-056-052-8

Query Match 72.3%; Score 47; DB 14; Length 118;  
Best Local Similarity 77.8%; Pred. No. 8.3;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 YYGNSWFAY 11  
||:|||||  
DB 99 YYGNSWFAY 107

RESULT 3  
US-10-056-052-16  
; Sequence 16, Application US/10056052  
; Publication No. US20030099656A1  
; GENERAL INFORMATION:  
; APPLICANT: PATTI, Joseph M  
; APPLICANT: HUTCHINS, Jeff T  
; APPLICANT: DOMANSKI, Paul  
; APPLICANT: PATEL, Pratiksha  
; APPLICANT: HALL, Andrea  
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN . . .  
; FILE REFERENCE: P07069US04/BAS  
; CURRENT APPLICATION NUMBER: US/10/056,052  
; CURRENT FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/308,116  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/298,413  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/274,611  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 60/264,072  
; PRIOR FILING DATE: 2001-01-26  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 118  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-056-052-16

Query Match 72.3%; Score 47; DB 14; Length 118;  
Best Local Similarity 77.8%; Pred. No. 8.3;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 YYGNSWFAY 11  
||:|||||

Db 99 YYGNSWFAY 107

RESULT 4  
US-10-417-895A-59  
; Sequence 59, Application US/10417895A  
; Publication No. US20040033569A1  
; GENERAL INFORMATION:  
; APPLICANT: Crea, Roberto  
; APPLICANT: Cappuccilli, Guido  
; TITLE OF INVENTION: "DOPING" IN WALK-THROUGH MUTAGENESIS  
; FILE REFERENCE: 1551.2002-001  
; CURRENT APPLICATION NUMBER: US/10/417,895A  
; CURRENT FILING DATE: 2003-04-16  
; PRIOR APPLICATION NUMBER: 60/373,686  
; PRIOR FILING DATE: 2002-04-17  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 59  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: variant peptide for third complementarity  
; OTHER INFORMATION: determining region of Fv region of an  
; OTHER INFORMATION: immunoglobulin  
US-10-417-895A-59

Query Match 69.2%; Score 45; DB 15; Length 11;  
Best Local Similarity 66.7%; Pred. No. 1.9;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 YYASSWFAY 11  
||:|||||  
DB 2 YYSWSWY 10

RESULT 5  
US-10-417-895A-57  
; Sequence 57, Application US/10417895A  
; Publication No. US20040033569A1  
; GENERAL INFORMATION:  
; APPLICANT: Crea, Roberto  
; APPLICANT: Cappuccilli, Guido  
; TITLE OF INVENTION: "DOPING" IN WALK-THROUGH MUTAGENESIS  
; FILE REFERENCE: 1551.2002-001  
; CURRENT APPLICATION NUMBER: US/10/417,895A  
; CURRENT FILING DATE: 2003-04-16  
; PRIOR APPLICATION NUMBER: 60/373,686  
; PRIOR FILING DATE: 2002-04-17  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 57  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: variant peptide for third complementarity  
; OTHER INFORMATION: determining region of Fv region of an  
; OTHER INFORMATION: immunoglobulin  
US-10-417-895A-57

Query Match 64.6%; Score 42; DB 15; Length 11;  
Best Local Similarity 66.7%; Pred. No. 5.6;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 YYASSWFAY 11  
||:|||||  
DB 2 YYSWSWY 10

RESULT 6  
US-10-435-614-11

```
; Sequence 11, Application US/10435614
; Publication No. US20040005647A1
; GENERAL INFORMATION:
; APPLICANT: DENARDO, SALLY J.
; APPLICANT: WINTHROP, MICHELLE D.
; APPLICANT: DENARDO, GERALD L.
; TITLE OF INVENTION: ANTI-MUC-1 SINGLE CHAIN ANTIBODIES FOR TUMOR TARGETING
; FILE REFERENCE: 309T-000220US
; CURRENT APPLICATION NUMBER: US/10/435,614
; CURRENT FILING DATE: 2003-05-09
; PRIOR FILING DATE: 2002-03-28
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 11
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Mus Musculus
US-10-435-614-11

Query Match      63.8%; Score 41.5; DB 15; Length 118;
Best Local Similarity 53.3%; Pred. No. 59;
Matches 8; Conservative 2; Mismatches 0; Indels 5; Gaps 1;

Qy  2 IYY-----ASSWFAY 11
Db  93 VYYCARDGSSSWFAY 107

RESULT 7
US-10-435-614-13
; Sequence 13, Application US/10435614
; Publication No. US20040005647A1
; GENERAL INFORMATION:
; APPLICANT: DENARDO, SALLY J.
; APPLICANT: WINTHROP, MICHELLE D.
; APPLICANT: DENARDO, GERALD L.
; TITLE OF INVENTION: ANTI-MUC-1 SINGLE CHAIN ANTIBODIES FOR TUMOR TARGETING
; FILE REFERENCE: 309T-000220US
; CURRENT APPLICATION NUMBER: US/10/435,614
; CURRENT FILING DATE: 2003-05-09
; PRIOR FILING DATE: 2002-03-28
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 13
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Mus Musculus
US-10-435-614-13

Query Match      63.8%; Score 41.5; DB 15; Length 118;
Best Local Similarity 53.3%; Pred. No. 59;
Matches 8; Conservative 2; Mismatches 0; Indels 5; Gaps 1;

Qy  2 IYY-----ASSWFAY 11
Db  93 VYYCARDGSSSWFAY 107

RESULT 8
US-10-469-125-39
; Sequence 39, Application US/10469125
; Publication No. US20040143101A1
; GENERAL INFORMATION:
; APPLICANT: Soltis, Daniel A.
; APPLICANT: Burch, Ronald M.
; APPLICANT: Shukla, Rajiv
; TITLE OF INVENTION: IMMUNOGLOBULIN CONSTRUCT CONTAINING ANTI-MUCIN VARIABLE DOMAIN SE
```

```
; TITLE OF INVENTION: FOR ELICITING AN ANTI-IDIOTYPE ANTI-TUMOR RESPONSE
; FILE REFERENCE: 02755/100G273-US1
; CURRENT APPLICATION NUMBER: US/10/469,125
; CURRENT FILING DATE: 2003-08-25
; PRIOR APPLICATION NUMBER: US 60/281,182
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: PCT/US02/10304
; PRIOR FILING DATE: 2002-04-02
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 39
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide (HMPG-1 heavy chain variable region)
US-10-469-125-39

Query Match      63.8%; Score 41.5; DB 16; Length 137;
Best Local Similarity 53.3%; Pred. No. 68;
Matches 8; Conservative 2; Mismatches 0; Indels 5; Gaps 1;

Qy  2 IYVAS-----SWFAY 11
Db  112 VYVASRSYDFAWFAY 126

RESULT 9
US-10-435-614-21
; Sequence 21, Application US/10435614
; Publication No. US20040005647A1
; GENERAL INFORMATION:
; APPLICANT: DENARDO, SALLY J.
; APPLICANT: WINTHROP, MICHELLE D.
; APPLICANT: DENARDO, GERALD L.
; TITLE OF INVENTION: ANTI-MUC-1 SINGLE CHAIN ANTIBODIES FOR TUMOR TARGETING
; FILE REFERENCE: 309T-000220US
; CURRENT APPLICATION NUMBER: US/10/435,614
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: 10/112,788
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,721
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 21
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Mus Musculus
US-10-435-614-21

Query Match      63.8%; Score 41.5; DB 15; Length 177;
Best Local Similarity 53.3%; Pred. No. 86;
Matches 8; Conservative 2; Mismatches 0; Indels 5; Gaps 1;

Qy  2 IYY-----ASSWFAY 11
Db  95 VYYCARDGSSSWFAY 109

RESULT 10
US-10-435-614-20
; Sequence 20, Application US/10435614
; Publication No. US20040005647A1
; GENERAL INFORMATION:
; APPLICANT: DENARDO, SALLY J.
; APPLICANT: WINTHROP, MICHELLE D.
; APPLICANT: DENARDO, GERALD L.
; TITLE OF INVENTION: ANTI-MUC-1 SINGLE CHAIN ANTIBODIES FOR TUMOR TARGETING
; FILE REFERENCE: 309T-000220US
; CURRENT APPLICATION NUMBER: US/10/435,614
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: 10/112,788
```

; PRIOR FILING DATE: 2002-03-28  
; PRIOR APPLICATION NUMBER: 60/280,721  
; PRIOR FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 20  
; LENGTH: 260  
; TYPE: PRT  
; ORGANISM: Mus Musculus  
US-10-435-614-20

Query Match 63.8%; Score 41.5; DB 15; Length 260;  
Best Local Similarity 53.3%; Pred. No. 1.2e+02;  
Matches 8; Conservative 2; Mismatches 0; Indels 5; Gaps 1;

Qy 2 IYV-----ASSWFAY 11  
|:|:|:|:|:  
Db 95 VYICARDGSSSWFAY 109

RESULT 11  
US-10-282-122A-72844  
; Sequence 72844, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zykkind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 72844  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: Salmonella paratyphi A  
US-10-282-122A-72844

Query Match 63.1%; Score 41; DB 15; Length 219;  
Best Local Similarity 55.6%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYASSWF 9  
|:|:|:|:  
Db 69 PLYYGSEWY 77

RESULT 12  
US-10-282-122A-75074  
; Sequence 75074, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zykkind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 75074  
; LENGTH: 232  
; TYPE: PRT  
; ORGANISM: Salmonella typhimurium  
US-10-282-122A-75074

Query Match 63.1%; Score 41; DB 15; Length 232;  
Best Local Similarity 55.6%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYASSWF 9  
|:|:|:|:  
Db 69 PLYYGSEWY 77

RESULT 13  
US-10-282-122A-75821  
; Sequence 75821, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl

```
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75821
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-10-282-122A-75821

Query Match      63.1%; Score 41; DB 15; Length 232;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 PYYVASSWF 9
Db      69 PLYYGSEWY 77
      |::|||::|
      |::|||::|

RESULT 14
US-10-282-122A-43712
; Sequence 43712, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms.
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
```

```
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43712
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-10-282-122A-43712

Query Match      61.5%; Score 40; DB 15; Length 125;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      4 YASSWFAY 11
Db      111 FASTWFAY 118
      |::|||::|
      |::|||::|

RESULT 15
US-10-425-114-47000
; Sequence 47000, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47000
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700548408_FLI.pep
US-10-425-114-47000

Query Match      61.5%; Score 40; DB 15; Length 206;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 IYYASSWFAY 11
Db      113 IYFASSWTF 122
      |::|||::|
      |::|||::|

RESULT 16
US-10-156-761-14201
```

```
; Sequence 14201, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14201
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14201

Query Match 61.5%; Score 40; DB 14; Length 376;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYVASSWF 9
Db 8 PVIYVAPSWF 16

RESULT 17
US-10-087-887-87
; Sequence 87, Application US/10087887
; Publication No. US20030198957A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Conley, Pamela B.
; APPLICANT: Yang, Ruey-Bing
; APPLICANT: Hart, Matthew
; APPLICANT: Tomlinson, James E.
; APPLICANT: Topper, James N.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Leach, Martin D.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Komuves, Laszlo
; APPLICANT: Padigar, Muralidhara
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-285
; CURRENT APPLICATION NUMBER: US/10/087,887
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/273,049
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/279,883
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/277,791
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/281,248
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/282,864
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,537
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/282,867
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 87
; LENGTH: 2174
; TYPE: PRT

; ORGANISM: Drosophila melanogaster
US-10-087-887-87

Query Match 61.5%; Score 40; DB 14; Length 2174;
Best Local Similarity 45.5%; Pred. No. 1.5e+03;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYVASSWFAY 11
Db 276 PMFFANSWNY 286

RESULT 18
US-09-927-668-2
; Sequence 2, Application US/09927668
; Patent No. US20020115144A1
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; APPLICANT: Moore, Charles
; APPLICANT: Perou, Charles
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/927,668
; FILING DATE: 10-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/396,540
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2186 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-927-668-2

Query Match 61.5%; Score 40; DB 9; Length 2186;
Best Local Similarity 75.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYVASSW 8
Db 1325 PIYPTSW 1332

RESULT 19
US-09-927-668-12
; Sequence 12, Application US/09927668
; Patent No. US20020115144A1
```



```
;
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; Perou, Charles
; Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/927,668
; FILING DATE: 10-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/396,540
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3672 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-927-668-12

Query Match 61.5%; Score 40; DB 9; Length 3672;
Best Local Similarity 75.0%; Pred. No. 2.4e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYVASSW 8
Db 2811 PIYPTSW 2818

RESULT 20
US-09-927-668-10
; Sequence 10, Application US/09927668
; Patent No. US20020115144A1
; GENERAL INFORMATION:
; APPLICANT: Kaplan, Jerry
; Perou, Charles
; Moore, Karen
; TITLE OF INVENTION: COMPOSITIONS FOR THE DIAGNOSIS
; AND TREATMENT OF CHEDIAK-HIGASHI SYNDROME
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
;
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/927,668
; FILING DATE: 10-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/396,540
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-062-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3801 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-927-668-10

Query Match 61.5%; Score 40; DB 9; Length 3801;
Best Local Similarity 75.0%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYVASSW 8
Db 2940 PIYPTSW 2947

RESULT 21
US-10-755-889-777
; Sequence 777, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 777
; LENGTH: 3801
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-755-889-777

Query Match 61.5%; Score 40; DB 16; Length 3801;
Best Local Similarity 75.0%; Pred. No. 2.5e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYVASSW 8
Db 2940 PIYPTSW 2947

RESULT 22
US-09-791-551-97
; Sequence 97, Application US/09791551
; Publication No. US20030235584A1
; GENERAL INFORMATION:
; APPLICANT: KLOETZER, WILLIAM S.
```

; APPLICANT: HANNA, NABIL  
; TITLE OF INVENTION: METHOD FOR PREPARING ANTI-MIF ANTIBODIES  
; FILE REFERENCE: 037003/0277869  
; CURRENT APPLICATION NUMBER: US/09/791,551  
; CURRENT FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 60/185,390  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/233,625  
; PRIOR FILING DATE: 2000-09-18  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 97  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-09-791-551-97

Query Match 60.8%; Score 39.5; DB 10; Length 14;  
Best Local Similarity 80.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 3 YYASS-WFAY 11  
||| |||  
Db 5 YGSSPWFA 14

## RESULT 23

US-09-791-551-111  
; Sequence 111, Application US/09791551  
; Publication No. US20030235584A1  
; GENERAL INFORMATION:  
; APPLICANT: HANNA, NABIL  
; TITLE OF INVENTION: METHOD FOR PREPARING ANTI-MIF ANTIBODIES  
; FILE REFERENCE: 037003/0277869  
; CURRENT APPLICATION NUMBER: US/09/791,551  
; CURRENT FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 60/185,390  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/233,625  
; PRIOR FILING DATE: 2000-09-18  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 111  
; LENGTH: 146  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-09-791-551-111

Query Match 60.8%; Score 39.5; DB 10; Length 146;  
Best Local Similarity 80.0%; Pred. No. 1.5e+02;  
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 3 YYASS-WFAY 11  
||| |||  
Db 124 YGSSPWFA 133

## RESULT 24

US-10-424-599-254210  
; Sequence 254210, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 254210  
; LENGTH: 81  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(81)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_71576C.1.pep  
US-10-424-599-254210

Query Match 60.0%; Score 39; DB 15; Length 81;  
Best Local Similarity 75.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IYYASSWF 9  
||| |||  
Db 34 IYYGSGWF 41

## RESULT 25

US-10-437-963-140199  
; Sequence 140199, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 140199  
; LENGTH: 118  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_4141C.1.pep  
US-10-437-963-140199

Query Match 60.0%; Score 39; DB 16; Length 118;  
Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASSWF 9  
||: ||: |||  
Db 95 PLYPSSSWF 103

## RESULT 26

US-10-887-231-21  
; Sequence 21, Application US/10887231  
; Publication No. US20050064512A1  
; GENERAL INFORMATION:  
; APPLICANT: Schering AG  
; APPLICANT: MorphoSys AG  
; TITLE OF INVENTION: Emitter-binding peptides that produce a change in the spectral  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/887,231  
; CURRENT FILING DATE: 2004-07-09  
; PRIOR APPLICATION NUMBER: DE 103 31 054.1  
; PRIOR FILING DATE: 2003-07-09  
; PRIOR APPLICATION NUMBER: US 60/487,234  
; PRIOR FILING DATE: 2003-07-16

; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 21  
; LENGTH: 216  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Fab fragment MOR03201 VL-CL  
US-10-887-231-21

Query Match 60.0%; Score 39; DB 17; Length 216;  
Best Local Similarity 66.7%; Pred. No. 2.5e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 YYASSWFAY 11  
||| : ||  
Db 88 YYCSWTSY 96

## RESULT 27

US-10-984-449-30  
; Sequence 30, Application US/10984449  
; Publication No. US20050089973A1  
; GENERAL INFORMATION:  
; APPLICANT: Yocum, R. Rogers  
; APPLICANT: Patterson, Thomas A.  
; APPLICANT: Hermann, Theron  
; APPLICANT: Pero, Janice G.  
; TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR PRODUCTION OF  
; TITLE OF INVENTION: PANTO-COMPOUNDS  
; FILE REFERENCE: BGI-141CPN  
; CURRENT APPLICATION NUMBER: US/10/984,449  
; CURRENT FILING DATE: 2004-11-08  
; PRIOR APPLICATION NUMBER: US9N 09/667,569  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US9N 09/400,494  
; PRIOR FILING DATE: 1999-09-21  
; PRIOR APPLICATION NUMBER: US9N 60/210,072  
; PRIOR FILING DATE: 2000-06-07  
; PRIOR APPLICATION NUMBER: US9N 60/221,836  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: US9N 60/227,860  
; PRIOR FILING DATE: 2000-08-24  
; NUMBER OF SEQ ID NOS: 94  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 30  
; LENGTH: 298  
; TYPE: PRT  
; ORGANISM: Bacillus subtilis  
US-10-984-449-30

Query Match 60.0%; Score 39; DB 17; Length 298;  
Best Local Similarity 55.6%; Pred. No. 3.4e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYYASSWF 9  
||| : ||  
Db 165 PIYYETDWF 173

## RESULT 28

US-10-369-493-6525  
; Sequence 6525, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 6525  
; LENGTH: 451  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-10-369-493-6525

Query Match 60.0%; Score 39; DB 15; Length 451;  
Best Local Similarity 63.6%; Pred. No. 5e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYYASSWFAY 11  
||| : |||  
Db 180 PIMYSCSWPAY 190

## RESULT 29

US-10-369-493-11837  
; Sequence 11837, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 11837  
; LENGTH: 520  
; TYPE: PRT  
; ORGANISM: Mesorhizobium loti  
US-10-369-493-11837

Query Match 60.0%; Score 39; DB 15; Length 520;  
Best Local Similarity 63.6%; Pred. No. 5.6e+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYYASSWFAY 11  
||| : |||  
Db 126 PIFRASDWEAY 136

## RESULT 30

US-10-437-963-185152  
; Sequence 185152, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 185152

```
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(727)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_82076C.1.pap
US-10-437-963-185152

Query Match          60.0%; Score 39; DB 16; Length 727;
Best Local Similarity 36.4%; Pred. No. 7.7e+02;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIYVASSWFAY 11
   |:|:|:|:|:|
Db 493 PVYFKNDWFPF 503

RESULT 31
US-10-101-464A-942
; Sequence 942, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR FILING DATE: 2002-03-18
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 942
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-942

Query Match          60.0%; Score 39; DB 14; Length 744;
Best Local Similarity 75.0%; Pred. No. 7.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 YASSWFAY 11
   |:|:|:|:|
Db 242 YINSWFAY 249

RESULT 32
US-10-864-252-942
; Sequence 942, Application US/10864252
; Publication No. US2005005083A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c3
; CURRENT APPLICATION NUMBER: US/10/864,252
; CURRENT FILING DATE: 2004-06-09
; PRIOR FILING DATE: 2002-03-18
```

```
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 942
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-864-252-942

Query Match          60.0%; Score 39; DB 17; Length 744;
Best Local Similarity 75.0%; Pred. No. 7.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 YASSWFAY 11
   |:|:|:|:|
Db 242 YINSWFAY 249

RESULT 33
US-10-032-585-7858
; Sequence 7858, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7858
; LENGTH: 1262
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7858

Query Match          60.0%; Score 39; DB 14; Length 1262;
Best Local Similarity 55.6%; Pred. No. 1.3e+03;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PIYVASSWF 9
   |:|:|:|:|
Db 218 PMFYAPNWF 226

RESULT 34
US-10-308-817-136
; Sequence 136, Application US/10308817
; Publication No. US20030219861A1
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; APPLICANT: Wu, Dayang
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: US/10/308,817
; CURRENT FILING DATE: 2002-12-03
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 136
; LENGTH: 115
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
```

; OTHER INFORMATION: hybrid humanized antibody heavy chain  
US-10-308-817-136

Query Match 59.2%; Score 38.5; DB 15; Length 115;  
Best Local Similarity 53.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

Qy 2 IYVAS---SWFAY 11  
: || : |||||  
Db 92 VYICARWESWFAY 104

## RESULT 35

US-10-453-698-136  
; Sequence 136, Application US/10453698  
; Publication No. US20040038308A1

; GENERAL INFORMATION:

; APPLICANT: Rother, Russell

; TITLE OF INVENTION: HYBRID ANTIBODIES

; FILE REFERENCE: 82 CIP (1087-37 CIP)

; CURRENT APPLICATION NUMBER: US/10/453,698

; CURRENT FILING DATE: 2003-06-03

; NUMBER OF SEQ ID NOS: 196

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 136

; LENGTH: 115

; TYPE: PRT

; ORGANISM: artificial sequence

; FEATURE:

; OTHER INFORMATION: hybrid humanized antibody heavy chain

US-10-453-698-136

Query Match 59.2%; Score 38.5; DB 15; Length 115;  
Best Local Similarity 53.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

Qy 2 IYVAS---SWFAY 11  
: || : |||||  
Db 92 VYICARWESWFAY 104

## RESULT 36

US-10-308-817-157

; Sequence 157, Application US/10308817

; Publication No. US20030219861A1

; GENERAL INFORMATION:

; APPLICANT: Rother, Russell

; APPLICANT: Wu, Dayang

; TITLE OF INVENTION: HYBRID ANTIBODIES

; FILE REFERENCE: 1087-37

; CURRENT APPLICATION NUMBER: US/10/308,817

; CURRENT FILING DATE: 2002-12-03

; NUMBER OF SEQ ID NOS: 195

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 157

; LENGTH: 116

; TYPE: PRT

; ORGANISM: artificial sequence

; FEATURE:

; OTHER INFORMATION: hybrid antibody variable heavy chain

US-10-308-817-157

Query Match 59.2%; Score 38.5; DB 15; Length 116;  
Best Local Similarity 53.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

Qy 2 IYVAS---SWFAY 11  
: || : |||||  
Db 93 VYICARWESWFAY 105

## RESULT 37

US-10-453-698-157

; Sequence 157, Application US/10453698  
; Publication No. US20040038308A1

; GENERAL INFORMATION:

; APPLICANT: Rother, Russell

; TITLE OF INVENTION: HYBRID ANTIBODIES

; FILE REFERENCE: 82 CIP (1087-37 CIP)

; CURRENT APPLICATION NUMBER: US/10/453,698

; CURRENT FILING DATE: 2003-06-03

; NUMBER OF SEQ ID NOS: 196

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 157

; LENGTH: 116

; TYPE: PRT

; ORGANISM: artificial sequence

; FEATURE:

; OTHER INFORMATION: hybrid antibody variable heavy chain

US-10-453-698-157

Query Match 59.2%; Score 38.5; DB 15; Length 116;  
Best Local Similarity 53.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

Qy 2 IYVAS---SWFAY 11  
: || : |||||  
Db 93 VYICARWESWFAY 105

## RESULT 38

US-10-437-963-130365

; Sequence 130365, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 130365

; LENGTH: 34

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(34)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_32534C.1.pep

US-10-437-963-130365

Query Match 58.5%; Score 38; DB 16; Length 34;  
Best Local Similarity 75.0%; Pred. No. 66;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 YASSWFAY 11  
: || : |||||  
Db 13 YRESWFAY 20

## RESULT 39

US-10-424-599-283720

; Sequence 283720, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 283720  
; LENGTH: 74  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_98222C.1.pep  
US-10-424-599-283720

Query Match 58.5%; Score 38; DB 15; Length 74;  
Best Local Similarity 45.5%; Pred. No. 1.3e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIYVASSWFAV 11  
||| : |||  
Db 36 PIFYGNPWAF 46

RESULT 40  
US-10-424-599-234514  
; Sequence 234514, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 234514  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(76)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_53794C.1.pep  
US-10-424-599-234514

Query Match 58.5%; Score 38; DB 15; Length 76;  
Best Local Similarity 62.5%; Pred. No. 1.4e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 IYVASSWF 9  
||| : |||  
Db 59 IYISTNWF 66

Search completed: June 3, 2005, 16:24:42  
Job time : 80.6667 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 15:48:25 ; Search time 19.3333 Seconds  
(without alignments)  
54.744 Million cell updates/sec

Title: US-10-089-452-23

Perfect score: 65

Sequence: 1 PIYVSSWFAY 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	67.7	107	2	S14506
2	43	66.2	91	2	S59647
3	42	64.6	26	2	PH1703
4	41	63.1	26	2	PH1718
5	41	63.1	26	2	PH1702
6	41	63.1	170	2	A35944
7	41	63.1	232	2	A80741
8	41	63.1	861	1	VCLJSC
9	40	61.5	142	2	C34903
10	40	61.5	389	2	F81211
11	40	61.5	389	2	A81787
12	40	61.5	3788	2	T30851
13	40	61.5	3788	2	T13960
14	40	61.5	3796	2	T18514
15	39.5	60.8	115	2	S26470
16	39	60.0	107	2	S14492
17	39	60.0	107	2	S14491
18	39	60.0	107	2	S14491
19	39	60.0	111	2	PH0994
20	39	60.0	298	2	E69875
21	39	60.0	399	2	E86289
22	39	60.0	405	2	S45522
23	39	60.0	451	2	T24018
24	39	60.0	617	2	T22175
25	39	60.0	621	2	S73155
26	38	58.5	26	2	PH1687
27	38	58.5	67	2	T12860
28	38	58.5	113	1	AVMSB7
29	38	58.5	117	2	JC2269

30	38	58.5	131	2	A27472	Ig heavy chain pre
31	38	58.5	203	2	A11689	hypothetical prote
32	38	58.5	206	2	H83762	hypothetical prote
33	38	58.5	319	2	T38533	sur1 protein homol
34	38	58.5	347	2	T23944	hypothetical prote
35	38	58.5	450	2	J01614	gastrin receptor -
36	38	58.5	874	2	AC2287	hypothetical prote
37	37.5	57.7	88	2	G25155	Ig heavy chain v r
38	37.5	57.7	118	2	C30560	Ig heavy chain v r
39	37	56.9	24	2	PH1697	Ig heavy chain v r
40	37	56.9	49	2	PL0097	Ig heavy chain v r
41	37	56.9	113	2	S25044	Ig heavy chain v r
42	37	56.9	119	2	S20640	Ig heavy chain v r
43	37	56.9	123	2	S24691	Ig heavy chain V6
44	37	56.9	157	2	F82941	hypothetical UU033
45	37	56.9	193	2	T06972	hypothetical prote

## ALIGNMENTS

## RESULT 1

S14506

Ig heavy chain V region - mouse

C:Species: Mus musculus (house mouse)

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Jul-1999

C:Accession: S14506

R:Chen, Q.; Stenzel-Poore, M.; Rittenberg, M.B.

submitted to the EMBL Data Library, March 1991

A:Description: Natural polyclonal antibodies differ from Ag-induced antibodies in VH (

A:Reference number: S14484

A:Accession: S14506

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-107 &lt;CHE&gt;

A:Cross-references: EMBL:X58651; NID:G51291; PIDN:CAA41508.1; PID:G51292

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotrimer; immunoglobulin

F:15-97/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 67.7%; Score 44; DB 2; Length 107;

Best Local Similarity 77.8%; Pred. No. 2.1;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 YYASSWFAY 11

Db 99 YYAYDWFAY 107

## RESULT 2

S59647

probable SNARE protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 14-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 09-Jul-2004

C:Accession: T38611; S59647

R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, July 1995

A:Reference number: Z21731

A:Accession: T38611

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-91 &lt;DE2&gt;

A:Cross-references: UNIPROT:Q09730; EMBL:Z50113; NID:G914878; PIDN:CAA90471.1; PID:G91488

A:Experimental source: strain 972h-; cosmid c31A2

C:Genetics:

A:Gene: SPAC31A2.13c

A:Map position: 1

A:introns: 39/3

Query Match

Best Local Similarity 66.2%; Score 43; DB 2; Length 91;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 IYYASSWF 9  
| | | | |  
Db 83 IYYASKWF 90

## RESULT 3

PH1703 Ig heavy chain V region (clone NP-7-16) - mouse (fragment)

C;Species: Mus musculus (house mouse)  
C;Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
C;Accession: PH1703  
R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
J. Exp. Med. 178, 295-307, 1993

A;Title: Antigen-driven B cell differentiation in vivo.

A;Reference number: PH1675; MUID:93301607; PMID:8315385

A;Accession: PH1703

A;Molecule type: mRNA

A;Residues: 1-26 <MCH>

A;Experimental source: B cell

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

Query Match 64.6%; Score 42; DB 2; Length 26;  
Best Local Similarity 63.6%; Pred. No. 1.1;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYYASSWFAY 11  
| | | | |  
Db 16 PYYGSSYFDY 26

## RESULT 4

PH1718

Ig heavy chain V region (clone ASC-16) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999

C;Accession: PH1718

R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
J. Exp. Med. 178, 295-307, 1993

A;Title: Antigen-driven B cell differentiation in vivo.

A;Reference number: PH1675; MUID:93301607; PMID:8315385

A;Accession: PH1718

A;Molecule type: mRNA

A;Residues: 1-26 <MCH>

A;Experimental source: B cell

A;Note: the authors translated the codon GTA for residue 11 as Thr and ACA for residue 1

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

Query Match 63.1%; Score 41; DB 2; Length 26;  
Best Local Similarity 63.6%; Pred. No. 1.7;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYYASSWFAY 11  
| | | | |  
Db 16 PYYGSSYFDY 26

## RESULT 5

PH1702

Ig heavy chain V region (clone NP-7-15) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999

C;Accession: PH1702

R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
J. Exp. Med. 178, 295-307, 1993

A;Title: Antigen-driven B cell differentiation in vivo.

A;Reference number: PH1675; MUID:93301607; PMID:8315385

A;Accession: PH1702

A;Molecule type: mRNA

A;Residues: 1-26 <MCH>

A;Experimental source: B cell

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

Query Match 63.1%; Score 41; DB 2; Length 26;  
Best Local Similarity 63.6%; Pred. No. 1.7;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYYASSWFAY 11  
| | | | |  
Db 16 PYYGSSYFDY 26

## RESULT 6

A35944

Ig gamma-2a chain - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 30-Nov-1990 #sequence\_revision 30-Nov-1990 #text\_change 16-Jul-1999

C;Accession: A35944

R;Igarashi, T.; Sato, M.; Katsube, Y.; Takio, K.; Tanaka, T.; Nakanishi, M.; Arata, Y.  
Biochemistry 29, 5727-5733, 1990

A;Title: Structure of a mouse immunoglobulin G that lacks the entire C-H1 domain: protei

A;Reference number: A35944; MUID:90344795; PMID:2116900

A;Accession: A35944

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-170 <IGA>

C;Superfamily: immunoglobulin C region; immunoglobulin homology

F;56-120/Domain: immunoglobulin homology <IMM>

Query Match 63.1%; Score 41; DB 2; Length 170;  
Best Local Similarity 70.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 IYYASSWFAY 11  
| | | | |  
Db 16 IYHYHPWFAY 25

## RESULT 7

AB0741

exodeoxyribonuclease X (EC 3.1.11.-) [imported] - Salmonella enterica subsp. enterica ser

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: This species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C;Accession: AB0741

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AB0741

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-232 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD05629.1; PID:gl6503126; GSPDB:GN00176

C;Genetics:

A;Gene: STY2084

Query Match 63.1%; Score 41; DB 2; Length 232;  
Best Local Similarity 55.6%; Pred. No. 14;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYYASSWF 9  
| | | | |  
Db 69 PLYGSEWY 77

## RESULT 8

VCLJSC

env polyprotein precursor - human immunodeficiency virus type 1 (isolate SC)

N;Alternate names: coat polyprotein

C;Species: human immunodeficiency virus type 1, HIV-1



A;Note: host Homo sapiens (man)  
C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Feb-1997  
C;Accession: B28922  
R;Gurgo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Sta  
Virology 164, 531-536, 1988  
A;Title: Envelope sequences of two new United States HIV-1 isolates.  
A;Reference number: A28922; MUID:86219542; PMID:3369091  
A;Accession: B28922  
A;Molecule type: DNA  
A;Residues: 1-861 <GUR>  
C;Genetics:  
A;Gene: env  
C;Superfamily: type E retrovirus env polyprotein  
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote  
F;1-29/Domain: signal sequence #status predicted <SG>  
F;30-861/Product: env polyprotein #status predicted <EP>  
F;87,129,135,140,143,159,163,187,198,234,241,262,276,295,301,302,333,340,356,362,386,396  
Query Match 63.1%; Score 41; DB 1; Length 861;  
Best Local Similarity 60.0%; Pred. No. 50;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 PIYVASSWFA 10  
| | | : | | |  
Db 215 PIYCARWFA 224  
| | | : | | |  
RESULT 9  
C34903  
Ig heavy chain precursor V region (5-27) - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 27-Jul-1990 #sequence\_revision 27-Jul-1990 #text\_change 16-Aug-1996  
C;Accession: C34903  
R;Bedzyk, W.D.; Herron, J.N.; Edmundson, A.B.; Voss Jr., E.W.  
J. Biol. Chem. 265, 133-138, 1990  
A;Title: Active site structure and antigen binding properties of idiotypically cross-rea  
A;Reference number: A34903; MUID:90094387; PMID:2104617  
A;Accession: C34903  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-142 <BED>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;34-115/Domain: immunoglobulin homology <IMW>  
Query Match 61.5%; Score 40; DB 2; Length 142;  
Best Local Similarity 66.7%; Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 3 YYASSWFA 11  
| | | : | | |  
Db 123 YYGYWFA 131  
| | | : | | |  
RESULT 10  
F81211  
Conserved hypothetical protein NMB0339 [imported] - Neisseria meningitidis (strain MC58  
C;Species: Neisseria meningitidis  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C;Accession: F81211  
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000  
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A;Reference number: A81000; MUID:20175755; PMID:10710307  
A;Accession: F81211  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-389 <TET>  
A;Cross-references: UNIPROT:Q9K148; GB:AE002390; NID:g7225550; PMID:AAF4078  
A;Experimental source: serogroup B, strain MC58

C;Genetics:  
A;Gene: NMB0339  
Query Match 61.5%; Score 40; DB 2; Length 389;  
Best Local Similarity 50.0%; Pred. No. 34;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 PIYVASSWFA 10  
| | | : | | |  
Db 16 PVFFAMGWFA 25  
| | | : | | |  
RESULT 11  
A81787  
probable periplasmic hypothetical protein NMA2148 [imported] - Neisseria meningitidis (s  
C;Species: Neisseria meningitidis  
C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C;Accession: A81787  
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jogle, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A;Reference number: A81775; MUID:20222556; PMID:10761919  
A;Accession: A81787  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-389 <PAR>  
A;Cross-references: UNIPROT:Q9JST2; GB:AL162758; NID:g7380672; PIDN:CAB8536  
A;Experimental source: serogroup A, strain Z2491  
C;Genetics:  
A;Gene: NMA2148  
Query Match 61.5%; Score 40; DB 2; Length 389;  
Best Local Similarity 50.0%; Pred. No. 34;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 PIYVASSWFA 10  
| | | : | | |  
Db 16 PVFFAMGWFA 25  
| | | : | | |  
RESULT 12  
T30851  
lysosomal trafficking regulator, long splice form - mouse  
N;Alternate names: beige protein homolog  
C;Species: Mus musculus (house mouse)  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T30851  
R;Barbosa, M.D.F.S.; Tchernev, V.T.; Kingmore, S.F.  
submitted to the EMBL Data Library, September 1996  
A;Description: Two bg or not two bg? Longest isoform of mouse Lyst (beige) gene.  
A;Reference number: Z20903  
A;Accession: T30851  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-3788 <BAR>  
A;Cross-references: UNIPROT:P97412; EMBL:U70015; NID:g1813541; PID:g1813542; PIDN:AAC530  
A;Experimental source: strain C57BL/6J  
C;Genetics:  
A;Gene: Lyst  
A;Map position: 1  
C;Keywords: alternative splicing  
Query Match 61.5%; Score 40; DB 2; Length 3788;  
Best Local Similarity 75.0%; Pred. No. 3.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 PIYVASSW 8  
| | | : | | |  
Db 2927 PIYPTSW 2934  
| | | : | | |  
RESULT 13

Tl3960  
beige protein homolog - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: Tl3960  
R;Mori, M.; Nishikawa, T.; Higuchi, K.; Nishimura, M.  
submitted to the EMBL Data Library, November 1998  
A;Description: Deletion in the beige gene of the beige rat due to recombination between  
A;Reference number: Z17837  
A;Accession: Tl3960  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-3788 <MOR>  
A;Cross-references: UNIPROT:Q922X9; EMBL:AB020019; NID:d1241953; PID:d1035670; PIDN:BAA3  
A;Experimental source: strain DA; spleen  
C;Genetics:  
A;Gene: beige

Query Match 61.5%; Score 40; DB 2; Length 3788;  
Best Local Similarity 75.0%; Pred. No. 3.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASSW 8  
|||:|  
Db 2927 PIYYPTSW 2934

RESULT 14  
Tl8514  
lysosomal trafficking regulator - bovine  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: Tl8514  
R;Yamakuchi, H.; Hirano, T.; Hara, K.; Sugimoto, Y.  
submitted to the EMBL Data Library, December 1998  
A;Description: Chediak-Higashi syndrome mutation and genetic testing in Japanese Black c  
A;Reference number: Z18939  
A;Accession: Tl8514  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-3796 <YAM>  
A;Cross-references: UNIPROT:O97699; EMBL:AF114785; NID:g4240557; PID:g4240558; PIDN:AADI  
C;Genetics:  
A;Gene: LYST

Query Match 61.5%; Score 40; DB 2; Length 3796;  
Best Local Similarity 75.0%; Pred. No. 3.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASSW 8  
|||:|  
Db 2935 PIYYPTSW 2942

RESULT 15  
S26470  
Ig heavy chain V region - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999  
C;Accession: S26470  
R;Kavaler, J.  
submitted to the EMBL Data Library, April 1991  
A;Reference number: S26459  
A;Accession: S26470  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-115 <KAV>  
A;Cross-references: EMBL:X59109; NID:g51948; PIDN:CAA41835.1; PID:g51949  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-97/Domain: immunoglobulin homology <IMM>

Query Match 60.8%; Score 39.5; DB 2; Length 115;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Best Local Similarity 80.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 3 YYASS-WFAY 11  
|||:|  
Db 100 YGSSXWFAY 109

RESULT 16  
S14492  
Ig heavy chain V region - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Jul-1999  
C;Accession: S14492  
R;Chen, Q.; Stenzel-Poore, M.; Rittenberg, M.B.  
submitted to the EMBL Data Library, March 1991  
A;Description: Natural polyclonal antibodies differ from Ag-induced antibodies in VH C  
A;Reference number: S14484  
A;Accession: S14492  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-107 <CHE>  
A;Cross-references: EMBL:X58649; NID:g51287; PIDN:CAA41506.1; PID:g51288  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-97/Domain: immunoglobulin homology <IMM>

Query Match 60.0%; Score 39; DB 2; Length 107;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 YYASSWFAY 11  
|||:|  
Db 99 YYRYDWFAY 107

RESULT 17  
S14493  
Ig heavy chain V region - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Jul-1999  
C;Accession: S14493  
R;Chen, Q.; Stenzel-Poore, M.; Rittenberg, M.B.  
submitted to the EMBL Data Library, March 1991  
A;Description: Natural polyclonal antibodies differ from Ag-induced antibodies in VH C  
A;Reference number: S14484  
A;Accession: S14493  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-107 <CHE>  
A;Cross-references: EMBL:X58650; NID:g51289; PIDN:CAA41507.1; PID:g51290  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-97/Domain: immunoglobulin homology <IMM>

Query Match 60.0%; Score 39; DB 2; Length 107;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 YYASSWFAY 11  
|||:|  
Db 99 YYRYDWFAY 107

RESULT 18  
S14491  
Ig heavy chain V region - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Jul-1999  
C;Accession: S14491  
R;Chen, Q.; Stenzel-Poore, M.; Rittenberg, M.B.  
submitted to the EMBL Data Library, March 1991  
A;Description: Natural polyclonal antibodies differ from Ag-induced antibodies in VH C

A:Reference number: S14484

A:Accession: S14491

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-107 <CHE>

A:Cross-references: EMBL:X58648; NID:g51285; PIDN:CAA41505.1; PID:g51286

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-97/Domain: immunoglobulin homology <IMM>

Query Match 60.0%; Score 39; DB 2; Length 107;

Best Local Similarity 66.7%; Pred. No. 14;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 YYASSWFAY 11

DB 99 YYRYDWFAY 107

RESULT 19

PH0994

Ig heavy chain V region (clone 202.135) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004

C:Accession: PH0994

E:Fillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.

J. Exp. Med. 176, 761-779, 1992

A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B c

A:Reference number: PH0971; MUID:92381444; PMID:1512540

A:Accession: PH0994

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-111 <TIU>

A:Cross-references: UNIPROT:Q9JL75

A:Experimental source: B cell, strain [NZB x NZW]F1

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 60.0%; Score 39; DB 2; Length 111;

Best Local Similarity 72.7%; Pred. No. 15;

Matches 8; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 3 YYAS--SWFAY 11

DB 100 YYGSSYSWFAY 110

RESULT 20

F69875

Pyrimidine-thiamin biosynthesis homolog ylbQ - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C:Accession: F69875

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Encian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y, M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,

A:Authors: Schlicht, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Banchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: F69875

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-298 <KUN>

A:Cross-references: UNIPROT:O34661; GB:Z99111; GB:AL009126; NID:g2633699; PIDN:CAB13384.

A:Experimental source: strain 168

C:Genetics:

A:Gene: ylbQ

Query Match 60.0%; Score 39; DB 2; Length 298;

Best Local Similarity 55.6%; Pred. No. 38;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PIYVASSWF 9

DB 165 PIYVETDWF 173

RESULT 21

E86289

Tl6N11.7 protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: E86289

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federgruel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Dewar, K.,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86289

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-399 <STO>

A:Cross-references: UNIPROT:Q9M9D8; GB:AE005172; NID:g8072393; PIDN:AAF71981.1; GSPDB:G1

C:Genetics:

A:Map position: 1

Query Match 60.0%; Score 39; DB 2; Length 399;

Best Local Similarity 50.0%; Pred. No. 51;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 PIYVASSWFA 10

DB 98 PIYKEAWFS 107

RESULT 22

S45522

alpha-N-acetylgalactosaminidase - chicken

C:Species: Gallus gallus (chicken)

C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004

C:Accession: S45522; S43413

R:Davis, M.O.; Hata, J.; Smith, D.; Walker, J.C.

submitted to the EMBL Data Library, December 1993

A:Reference number: S45522

A:Accession: S45522

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-405 <DAV>

A:Cross-references: UNIPROT:Q90744; EMBL:L18754; NID:g435036; PIDN:AAAL6614.1; PID:g4350

R:Davis, M.O.; Hata, D.J.; Smith, D.; Walker, J.C.

Biochim. Biophys. Acta 1216, 296-298, 1993

A:Title: Cloning and sequence of a chicken alpha-N-acetylgalactosaminidase gene.

A:Reference number: S43413; MUID:94060104; PMID:8241271

A:Accession: S43413

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-161, 'A', 163-405 <DA2>

A:Cross-references: GB:L18754

C:Superfamily: alpha-galactosidase

Query Match 60.0%; Score 39; DB 2; Length 405;  
 Best Local Similarity 63.6%; Pred. No. 51;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIYYASSWFAY 11  
 |||:|||||  
 Db 166 PIYSCSWPAY 176

RESULT 23  
 T24018  
 hypothetical protein R07B7.11 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T24018  
 R/Harris, B.  
 submitted to the EMBL Data Library, July 1996  
 A/Reference number: Z19830  
 A/Accession: T24018  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-451 <WIL>  
 A/Cross-references: UNIPROT:Q21801; EMBL:Z75955; PIDN:CAB00120.1; GSPDB:GN00023; CESP:RC0  
 A/Experimental source: clone R07B7  
 C/Genetics:  
 A/Gene: CESP:R07B7.11  
 A/Map position: 5  
 A/Introns: 61/1; 107/3; 317/3; 351/3  
 C/Superfamily: alpha-galactosidase

Query Match 60.0%; Score 39; DB 2; Length 451;  
 Best Local Similarity 63.6%; Pred. No. 57;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIYYASSWFAY 11  
 |||:|||||  
 Db 180 PIYSCSWPAY 190

RESULT 24  
 T22175  
 hypothetical protein F44F1.1 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T22175  
 R/Dobson, R.  
 submitted to the EMBL Data Library, October 1996  
 A/Reference number: Z19527  
 A/Accession: T22175  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-617 <WIL>  
 A/Cross-references: UNIPROT:O02259; EMBL:Z81083; PIDN:CAB03099.1; GSPDB:GN00019; CESP:F4  
 A/Experimental source: clone F44F1  
 C/Genetics:  
 A/Gene: CESP:F44F1.1  
 A/Map position: 1  
 A/Introns: 63/2; 79/2; 167/3; 382/3; 479/3; 512/2; 542/3

Query Match 60.0%; Score 39; DB 2; Length 617;  
 Best Local Similarity 75.0%; Pred. No. 77;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 YASSWFAY 11  
 |||:|||||  
 Db 336 YRSNWFAY 343

RESULT 25  
 S73155  
 hypothetical protein 621 - red alga (Porphyra purpurea) chloroplast  
 C/Species: chloroplast Porphyra purpurea

C/Date: 19-Mar-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004  
 C/Accession: S73155  
 R/Raith, M.; Munholland, J.  
 Plant Mol. Biol. Rep. 13, 333-335, 1995  
 A/Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.  
 A/Reference number: S73108  
 A/Accession: S73155  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-621 <REI>  
 A/Cross-references: UNIPROT:P51234; EMBL:U38804; NID:g1276652; PID:g1276700  
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995  
 C/Genetics:  
 A/Genome: chloroplast  
 C/Keywords: chloroplast

Query Match 60.0%; Score 39; DB 2; Length 621;  
 Best Local Similarity 54.5%; Pred. No. 78;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIYVASSWFAY 11  
 |||:|||||  
 Db 192 PLYLANSWFAY 202

RESULT 26  
 PH1687  
 IG heavy chain V region (clone NP-6-18) - mouse (fragment)  
 C/Species: Mus musculus (house mouse)  
 C/Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
 C/Accession: PH1687  
 R/McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
 J. Exp. Med. 178, 295-307, 1993  
 A/Title: Antigen-driven B cell differentiation in vivo.  
 A/Reference number: PH1675; MUID:93301607; PMID:8315385  
 A/Accession: PH1687  
 A/Molecule type: mRNA  
 A/Residues: 1-26 <MCH>  
 A/Experimental source: B cell  
 C/Superfamily: immunoglobulin V region; immunoglobulin homology  
 C/Keywords: heterotetramer; immunoglobulin

Query Match 58.5%; Score 38; DB 2; Length 26;  
 Best Local Similarity 75.0%; Pred. No. 5.2;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 YASSWFAY 11  
 |||:|||||  
 Db 19 YDSANWFAY 26

RESULT 27  
 T12860  
 hypothetical protein yopZ - Bacillus subtilis phage SPBc2  
 C/Species: Bacillus subtilis phage SPBc2  
 C/Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
 C/Accession: T12860; H69918  
 R/Lazarevic, V.; Duesterhoef, A.; Soldo, B.; Hilbert, H.; Mael, C.; Karamata, D.  
 submitted to the EMBL Data Library, August 1997  
 A/Description: The complete nucleotide sequence of the Bacillus subtilis SPbetac2 prophage  
 A/Reference number: Z17583  
 A/Accession: T12860  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-67 <LAZ>  
 A/Cross-references: UNIPROT:O64109; EMBL:AF020713; NID:g3025478; PID:g3025574; PIDN:AAC1  
 R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero  
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Puma, S.; Galizzi, A.; Gallera  
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadia, Y.; Sato, T.; Scanlon, A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowaka, A.; Seron, A;Authors: Tanakoshi, A.; Tanaka, T.; Terpatra, P.; Togononi, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A;Reference number: A59580; MUID:98044033; PMID:938377  
A;Accession: H69918  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-67 <KUN>  
A;Cross-references: GB:Z99114; GB:Z99115; GB:AL009126; NID:G2634478; PIDN:CAB13989.1; PI  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: yopZ

Query Match 58.5%; Score 38; DB 2; Length 67;  
Best Local Similarity 50.0%; Pred. No. 13; Mismatches 0; Gaps 0;  
Matches 5; Conservative 3; Indels 2; Indels 0; Gaps 0;  
Oy 2 IYYASSWFAY 11  
Db 47 LFPASSMIVY 56  
:::|||||  
A;Title: Heavy-chain variable-region sequence from an inulin-binding myeloma protein.  
A;Reference number: A90400; MUID:77134726; PMID:402936  
A;Accession: A90400  
A;Molecule type: protein  
A;Residues: 1-113 <VRA>  
A;Cross-references: UNIPROT:P01799  
A;Comment: This chain was isolated from a myeloma protein that binds inulin.  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-100/Domain: immunoglobulin homology <IMM>  
F;22-98/Disulfide bonds: #status predicted

Query Match 58.5%; Score 38; DB 1; Length 113;  
Best Local Similarity 70.0%; Pred. No. 22; Mismatches 1; Indels 2; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Mismatches 0; Gaps 0;  
Oy 2 IYYASSWFAY 11  
Db 95 IYYCSTGFAY 104  
|||:|:|  
A;Title: Antibody heavy chain - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 07-May-1999  
C;Accession: J02289; PC2186  
R;Kurume, T.; Katayama, M.; Murakami, K.; Hashino, K.; Kamihagi, K.; Yasumoto, M.; Kato, J. Biochem. 115, 608-614, 1994  
A;Title: Expression of recombinant mouse/human chimeric antibody specific to human GMP-140.  
A;Reference number: J02289; MUID:94334310; PMID:7520038  
A;Accession: J02289  
A;Molecule type: mRNA  
A;Residues: 1-117 <KUR>  
A;Accession: PC2186  
A;Molecule type: protein  
A;Residues: 2-27 <KU2>  
A;Experimental source: hybridoma cell

C;Comment: This protein is specific to human P-selectin.  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
F;2-98/Region: V segment  
F;15-98/Domain: immunoglobulin homology <IMM>  
F;99-106/Region: D segment  
F;107-117/Region: J segment  
Query Match 58.5%; Score 38; DB 2; Length 117;  
Best Local Similarity 50.0%; Pred. No. 23; Mismatches 1; Indels 4; Gaps 1;  
Matches 7; Conservative 2; Mismatches 1; Indels 4; Gaps 1;  
Oy 2 IYYASS---WFAY 11  
Db 93 VYYCAGNPAWFAY 106  
:|:|:|:|  
A;Title: Expression of mouse:human immunoglobulin heavy-chain cDNA in lymphoid cells.  
A;Reference number: A27472; MUID:87277430; PMID:3111940  
A;Accession: A27472  
A;Molecule type: mRNA  
A;Residues: 1-131 <LIU>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;5-54/Region: complementarity-determining 1  
F;20-131/Product: Ig heavy chain V region Ie9 #status predicted <MAT>  
F;34-117/Domain: immunoglobulin homology <IMM>  
F;69-85/Region: complementarity-determining 2  
F;118-125/Region: complementarity-determining 3

Query Match 58.5%; Score 38; DB 2; Length 131;  
Best Local Similarity 50.0%; Pred. No. 25; Mismatches 2; Mismatches 1; Indels 4; Gaps 1;  
Matches 7; Conservative 2; Mismatches 1; Indels 4; Gaps 1;  
Oy 2 IYYASS---WFAY 11  
Db 112 VYYCASYDYDFAY 125  
:|:|:|:|  
A;Title: Heavy-chain variable-region sequence from an inulin-binding myeloma protein.  
A;Reference number: A90400; MUID:77134726; PMID:402936  
A;Accession: A90400  
A;Molecule type: protein  
A;Residues: 1-113 <VRA>  
A;Cross-references: UNIPROT:P01799  
A;Comment: This chain was isolated from a myeloma protein that binds inulin.  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-100/Domain: immunoglobulin homology <IMM>  
F;22-98/Disulfide bonds: #status predicted

Query Match 58.5%; Score 38; DB 2; Length 131;  
Best Local Similarity 50.0%; Pred. No. 25; Mismatches 2; Mismatches 1; Indels 4; Gaps 1;  
Matches 7; Conservative 2; Mismatches 1; Indels 4; Gaps 1;  
Oy 2 IYYASS---WFAY 11  
Db 112 VYYCASYDYDFAY 125  
:|:|:|:|  
A;Title: Heavy-chain variable-region sequence from an inulin-binding myeloma protein.  
A;Reference number: A90400; MUID:77134726; PMID:402936  
A;Accession: A90400  
A;Molecule type: protein  
A;Residues: 1-113 <VRA>  
A;Cross-references: UNIPROT:P01799  
A;Comment: This chain was isolated from a myeloma protein that binds inulin.  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-100/Domain: immunoglobulin homology <IMM>  
F;22-98/Disulfide bonds: #status predicted

Query Match 58.5%; Score 38; DB 2; Length 203;  
Best Local Similarity 66.7%; Pred. No. 38; Mismatches 0; Mismatches 3; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Oy 2 IYYASS---WFAY 11  
Db 112 VYYCASYDYDFAY 125  
:|:|:|:|

```

Qy 1 PIYYASSWF 9
    |||||
Db 95 PIYLFHWF 103

RESULT 32
H83762
hypothetical protein BH0904 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H83762
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H83762
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-206 <STO>
A;Cross-references: UNIPROT:Q9KEE8; GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA046
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0904
C;Superfamily: Bacillus subtilis conserved hypothetical protein yesL

Query Match 58.5%; Score 38; DB 2; Length 206;
Best Local Similarity 55.6%; Pred. No. 39;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YYASSWFAY 11
    :||:|||||
Db 134 HYOTTWFAY 142

RESULT 33
T38533
srl1 protein homolog - fission yeast (Schizosaccharomyces pombe)
N;Alternate names: hypothetical protein SPAC2F3.01; hypothetical protein SPAC323.09
C;Species: Schizosaccharomyces pombe
C;Date: 20-Oct-2000 #sequence_revision 03-Nov-2000 #text_change 09-Jul-2004
C;Accession: T38646; T38533
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Ramsperger, U.; Pohl, T.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21802
A;Accession: T38646
A;Molecule type: DNA
A;Residues: 1-37 <WOO>
A;Cross-references: UNIPROT:O14084; EMBL:AL109988; PIDN:CAB53412.1; GSPDB:GN000066; SPDB:
A;Experimental source: strain 972h(-); cosmid c323
R;Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21748
A;Accession: T38533
A;Molecule type: DNA
A;Residues: 22-319 <OLI>
A;Cross-references: EMBL:Z99165; PIDN:CAB16259.1; GSPDB:GN000066; SPDB:SPAC2F3.01
A;Experimental source: strain 972h(-); cosmid c2F3
C;Genetics:
A;Gene: SPDB:SPAC2F3.01; SPDB:SPAC323.09
A;Map position: 1

Query Match 58.5%; Score 38; DB 2; Length 319;
Best Local Similarity 60.0%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYYASSWFA 10
    |||||
Db 175 PIGYNDWFA 184

RESULT 34
T23944
hypothetical protein R05H5.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23944
R;McMurray, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: Z19821
A;Accession: T23944
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-347 <WIL>
A;Cross-references: UNIPROT:Q21767; EMBL:Z48795; PIDN:CAA88731.1; GSPDB:GN000020; CESP:R0
A;Experimental source: clone R05H5
C;Genetics:
A;Gene: CESP:R05H5.1
A;Map position: 2
A;Introns: 35/1; 81/2; 194/1; 232/3; 257/3; 316/2

Query Match 58.5%; Score 38; DB 2; Length 347;
Best Local Similarity 54.5%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 PIYYASSWFAY 11
    |||||
Db 96 PCYVSPWFCY 106

RESULT 35
JQ1614
gastrin receptor - multimammate rat (Mastomys natalensis)
C;Species: Mastomys natalensis
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: JQ1614
R;Nakata, H.; Matsui, T.; Ito, M.; Taniguchi, T.; Naribayashi, Y.; Arima, N.; Nakamura,
Biochem. Biophys. Res. Commun. 187, 1151-1157, 1992
A;Title: Cloning and characterization of gastrin receptor from ECL carcinoid tumor of Ma
A;Reference number: JQ1614; MUID:92412082; PMID:1530611
A;Accession: JQ1614
A;Molecule type: mRNA
A;Residues: 1-450 <NAK>
A;Cross-references: UNIPROT:P30796; GB:D12817; NID:g220646; PIDN:BAA02250.1; PID:g220647
C;Superfamily: neurokinin 1 receptor
C;Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane pr
F;59-83/Domain: transmembrane #status predicted <TM1>
F;87-109/Domain: transmembrane #status predicted <TM2>
F;132-150/Domain: transmembrane #status predicted <TM3>
F;172-188/Domain: transmembrane #status predicted <TM4>
F;216-243/Domain: transmembrane #status predicted <TM5>
F;334-357/Domain: transmembrane #status predicted <TM6>
F;380-398/Domain: transmembrane #status predicted <TM7>
F;7,30,36/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 58.5%; Score 38; DB 2; Length 450;
Best Local Similarity 54.5%; Pred. No. 83;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYYASSWFAY 11
    |||||
Db 351 PIYSANTWCAF 361

RESULT 36
AC2287
hypothetical protein all3850 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AC2287
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat
A;Reference number: AB1807; MUID:21595285; PMID:11759840

```

```
A:Accession: AC2287
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-874 <KUR>
A:Cross-references: UNIPROT:Q8YQH8; GB:BA000019; PIDN:BA075549.1; PID:g17132984; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all3850

Query Match      58.5%; Score 38; DB 2; Length 874;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 PIYVASSWFA 10
Db      704 PTTTASSWLA 713

RESULT 37
G25155
Ig heavy chain V region (D11) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Aug-1996
C:Accession: G25155
R:Robbins, P.F.; Rosen, E.M.; Haba, S.; Nisonoff, A.
Proc. Natl. Acad. Sci. U.S.A. 83, 1050-1054, 1986
A:Title: Relationship of V(H) and V(L) genes encoding three idiotypic families of anti-P
A:Reference number: A94083; MUID:86149212; PMID:3081888
A:Accession: G25155
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-88 <ROB>
C:Comment: This chain is expressed in an IgG with anti-arsonate activity.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match      57.7%; Score 37.5; DB 2; Length 88;
Best Local Similarity 63.6%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy      2 IYV-ASSWFA 11
Db      67 VYFCARSWFA 77

RESULT 38
C30560
Ig heavy chain V region (35.8.2H) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-1989 #sequence_revision 03-Aug-1992 #text_change 23-Jul-1999
C:Accession: C30560
R:Matsuda, T.; Kabat, E.A.
J. Immunol. 142, 863-870, 1989
A:Title: Variable region cDNA sequences and antigen binding specificity of mouse monoclon
A:Reference number: A30560; MUID:89110062; PMID:2464028
A:Accession: C30560
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-118 <MAT>
A:Cross-references: GB:M24270; NID:g195615; PIDN:AAA38371.1; PID:g195616
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match      57.7%; Score 37.5; DB 2; Length 118;
Best Local Similarity 46.7%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

Qy      2 IYVAS-----SWFA 11
Db      93 VYICARWGTGSWFA 107
```

## RESULT 39

PH1697

Ig heavy chain V region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999

C:Accession: PH1697; PH1699

R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.

J. Exp. Med. 178, 295-307, 1993

A:Title: Antigen-driven B cell differentiation in vivo.

A:Reference number: PH1675; MUID:93301607; PMID:8315385

A:Accession: PH1697

A:Molecule type: mRNA

A:Residues: 1-24 &lt;MCH&gt;

A:Experimental source: B cell; clone NP-7-10

A:Note: the authors translated the codon GCT for residue 23 as Asp

A:Accession: PH1699

A:Molecule type: mRNA

A:Residues: 1-10,'A',12-13,'R',15-20,'T',22,'D',24 &lt;MC2&gt;

A:Experimental source: B cell; clone NP-7-12

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

## Query Match

Best Local Similarity 56.9%; Score 37; DB 2; Length 24;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 YYASSWFA 11

Db 16 YYGSSLEFA 24

## RESULT 40

PL0097

Ig heavy chain V region (GB4-10) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 30-May-1997

C:Accession: PL0097

R:Meek, K.; Hasemann, C.; Pollok, B.; Alkan, S.S.; Brait, M.; Slaoui, M.; Urbain, J.; Ca

J. Exp. Med. 169, 519-533, 1989

A:Title: Structural characterization of antiidiotypic antibodies; evidence that Ab2s are

A:Reference number: PL0080; MUID:89094248; PMID:2492056

A:Accession: PL0097

A:Molecule type: mRNA

A:Residues: 1-49 &lt;MEE&gt;

A:Note: the sequence shown here is from the VH region an allogenic antibody to antiphosp

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

## Query Match

Best Local Similarity 56.9%; Score 37; DB 2; Length 49;

Matches 7; Conservative 2; Mismatches 1; Indels 6; Gaps 1;

Qy 2 IYVA-----SSWFA 11

Db 24 VYICAREGGYSNWFAY 39

Search completed: June 3, 2005, 16:02:34

Job time : 21.3333 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 15:40:50 ; Search time 93.3333 Seconds  
(without alignments)  
60.352 Million cell updates/sec

Title: US-10-089-452-23

Perfect score: 65

Sequence: 1 PIYVASSWFAY 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	70.8	558	2	Q6ACC9
2	43	66.2	91	1	YA4D_SCHPO
3	41	63.1	232	2	Q8XFN9
4	41	63.1	232	2	Q7CQC9
5	41	63.1	294	2	Q8PW95
6	41	63.1	856	1	ENV_HV1SC
7	40	61.5	120	2	Q920R8
8	40	61.5	153	2	Q8FCY4
9	40	61.5	340	2	Q8MDY6
10	40	61.5	364	2	Q7N7V2
11	40	61.5	374	2	Q9RJ56
12	40	61.5	376	2	Q828J6
13	40	61.5	389	2	Q8J5W2
14	40	61.5	389	2	Q9K148
15	40	61.5	465	2	Q6LJ85
16	40	61.5	519	2	Q6BYK0
17	40	61.5	535	2	Q747F4
18	40	61.5	2174	2	Q9GQR0
19	40	61.5	2772	2	Q9VAV4
20	40	61.5	2776	2	Q869A0
21	40	61.5	2894	2	Q7KRX2
22	40	61.5	2898	2	Q868Z9
23	40	61.5	3788	1	LYST_MOUSE
24	40	61.5	3788	2	Q922X9
25	40	61.5	3795	2	Q97699
26	40	61.5	3796	2	Q9TTK4
27	40	61.5	3801	1	LYST_HUMAN
28	39.5	60.8	879	2	Q6C293
29	39	60.0	190	2	Q6H422
30	39	60.0	298	1	PANE_BACSU
31	39	60.0	399	2	Q9M9D8

32 39 60.0 400 2 Q6GR44 Q6gr44 xenopus lae  
33 39 60.0 405 2 Q90744 Q90744 gall  
34 39 60.0 428 2 Q7Q6H3 Q7q6h3 anopheles g  
35 39 60.0 431 2 Q6GQ17 Q6gq17 xenopus lae  
36 39 60.0 451 2 Q21801 Q21801 caenorhabdi  
37 39 60.0 500 2 Q02259 Q02259 caenorhabdi  
38 39 60.0 520 2 Q98M81 Q98m81 rhizobium 1  
39 39 60.0 530 2 Q96ZHG Q96zh6 sulfolobus  
40 39 60.0 534 2 Q6C9N1 Q6c9n1 yarrowia li  
41 39 60.0 594 2 Q86K08 Q86k08 dictyosteli  
42 39 60.0 621 1 YCXB\_PORPU P51234 porphyra pu  
43 39 60.0 1057 2 Q7SHK7 Q7shk7 neurospora  
44 39 60.0 1376 2 Q97AU9 Q97au9 thermoplasma  
45 39 60.0 2013 2 Q8I5C0 Q8i5c0 plasmodium

#### ALIGNMENTS

##### RESULT 1

ID Q6ACC9 PRELIMINARY; PRT; 558 AA.  
AC Q6ACC9;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DE ABC transporter, substrate binding protein.  
GN Name=QppA; OrderedLocuNames=Lxx23050;  
OS Leifsonia xyl (subsp. xyl).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Micrococciaceae; Microbacteriaceae; Leifsonia.  
OX NCBI\_TaxID=59736;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CTCB07;  
RX PubMed=15305603;  
RA Montelero-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,  
RA Kitajima J.P., Truffi D., do Amaral A.M., Harakava R.,  
RA de Oliveira J.C.F., Wood D., de Oliveira M.C., Miyaki C.Y.,  
RA Takita M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G.,  
RA Almeida N.F. Jr., Carter H., Coutinho L.L., El-Dorzy H.A.,  
RA Ferro M.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S.,  
RA Goldman G.H., Kimura E.T., Ferro E.S., Kuramae E.E., Lemos E.G.M.,  
RA Lemos M.V.F., Mauro S.M.Z., Machado M.A., Marino C.L., Menck C.F.,  
RA Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,  
RA Teai S.M., Zanca A.S., Simpson A.J.G., Brumbley S.M., Setubal J.C.,  
RT "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia  
RT xyl subsp. xyl."  
RL Mol. Plant Microbe Interact. 17:827-836(2004).  
CC -!- SIMILARITY: Belongs to the bacterial extracellular solute-binding  
CC protein family 5.  
DR EMBL; AB016822; AAT89964.1; -.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR006162; Ppantne\_S.  
DR InterPro; IPR000914; SBP\_bac\_5.  
DR Pfam; PF00496; SBP\_bac\_5; 1.  
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 558 AA; 61368 MW; B6625DC1351339F2 CRC64;

Query Match 70.8%; Score 46; DB 2; Length 558;

Best Local Similarity 54.5%; Pred. No. 24;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYVASSWFAY 11

Db 516 PLWIGASWFY 526

##### RESULT 2

YA4D\_SCHPO

ID\_YA4D\_SCHPO STANDARD; PRT; 91 AA.

```
AC Q09730;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical protein C31A2.13c in chromosome I.
GN ORFNames=SPAC31A2.13c;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Spours J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Shipakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z50113; CAA90471.1; -.
DR PIR; T38611; S59647.
DR GeneDB Spombe; SPAC31A2.13c; -.
DR Hypothetical protein; Transmembrane.
KW TRANSMEM 68 84 Potential.
SQ SEQUENCE 91 AA; 10342 MW; 78F8695CE8B16CCB CRC64;
Query Match 66.2%; Score 43; DB 1; Length 91;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 IYVASSWF 9
Db 83 IYVASKWF 90
RESULT 3
Q8XFN9 PRELIMINARY; PRT; 232 AA.
AC Q8XFN9; Q7AMT5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 23-OCT-2004 (TrEMBLrel. 28, Last annotation update)
CC -----
```

```
DE Exodeoxyribonuclease X (EC 3.1.11.-).
GN Name=exoX; OrderedLocusNames=STY2084, t0999;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TY2 / ATCC 700931;
RC MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RC STRAIN=CT18;
RX MEDLINE=22534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL; AE016837; AAO68671.1; -.
DR EMBL; AL627272; CAD05629.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR Pfam; PF00929; Exonuc X-T; 1.
DR SMART; SM00479; EXOIII; 1.
KW Complete proteome.
SQ SEQUENCE 232 AA; 26419 MW; 8FF0AC491F3A5EAC CRC64;
Query Match 63.1%; Score 41; DB 2; Length 232;
Best Local Similarity 55.8%; Pred. No. 67;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 PIYVASSWF 9
Db 69 PLYYGSWF 77
RESULT 4
Q7CQC9 PRELIMINARY; PRT; 232 AA.
ID AC Q7CQC9;
ID Q7CQC9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE DNA exonuclease X (EC 3.1.11.-).
GN Name=exoX; OrderedLocusNames=STM1878;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RC MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
```

```

RL Nature 413:852-856(2001).
DR EMBL; AE008783; AAL20794.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR InterPro; IPR006055; Exonuclease activity; IEA.
DR Pfam; PF00929; Exonuc X-T; 1.
DR SMART; SM00479; EXOIII; 1.
KW Complete proteome.
SQ SEQUENCE 232 AA; 26419 MW; 8FF0AC491F3A5EAC CRC64;

Query Match 63.1%; Score 41; DB 2; Length 232;
Best Local Similarity 55.6%; Pred. No. 67;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYYASSWF 9
Db 69 PLYYGSEWY 77

RESULT 5
Q8PW95 PRELIMINARY; PRT; 294 AA.
AC Q8PW95;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative nucleoside-diphosphate-sugar epimerase.
GN OrderedLocusNames=MM1701;
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Gos1 / Gol / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppe-meier U, Johann A, Hartsch T, Merkl R, Schmitz R.A.,
RA Brueggemann H, Lienard T, Christmann A, Baumer S, Jacobi C.,
RA Bhattacharyya A, Lykidis A, Overbeek R, Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
RW EMBL; AE013406; AAM31397.1; -.
KW Complete proteome.
SQ SEQUENCE 294 AA; 32685 MW; 1501DDC24DA37574 CRC64;

Query Match 63.1%; Score 41; DB 2; Length 294;
Best Local Similarity 54.5%; Pred. No. 86;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PIYYASSWF 11
Db 148 PAYYFSNWLY 158

RESULT 6
ENV_HVISC STANDARD; PRT; 856 AA.
AC P05878;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN Name=ENV;
OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;

```

```

RA RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates.";
RL Virology 164:531-536(1988).
CC -I- MISCELLANEOUS: The SC isolate was taken from an ARC patient in
CC 1984 in Southern California.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M17450; -, NOT_ANNOTATED_CDS.
DR DR HSP; P04578; 1DLB.
DR HIV; M17450; ENVSSC.
DR InterPro; IPR011010; DNA_brk_join_enz.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GPI20.
DR Pfam; PF00516; GPI20; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Signal; Transmembrane.
FT SIGNAL 1 29
FT CHAIN 30 510 Exterior membrane glycoprotein.
FT CHAIN 511 856 Transmembrane glycoprotein.
FT SITE 760 760 In-frame termination codon.
FT DISULFID 53 73 By similarity.
FT DISULFID 118 206 By similarity.
FT DISULFID 125 197 By similarity.
FT DISULFID 130 160 By similarity.
FT DISULFID 219 247 By similarity.
FT DISULFID 228 239 By similarity.
FT DISULFID 296 330 By similarity.
FT DISULFID 376 439 By similarity.
FT DISULFID 383 412 By similarity.
FT CARBOHYD 87 87 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 129 129 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 130 135 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 140 140 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 143 143 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 159 159 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 163 163 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 189 189 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 198 198 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 234 234 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 241 241 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 262 262 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 276 276 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 295 295 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 301 301 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 302 302 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 331 331 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 338 338 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 354 354 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 360 360 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 384 384 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 394 394 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 400 400 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 405 405 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 442 442 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 457 457 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 611 611 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 616 616 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 625 625 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 637 637 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 674 674 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 816 816 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 856 AA; 97055 MW; DAF4DA600EBA7A08 CRC64;

Query Match 63.1%; Score 41; DB 1; Length 856;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;

```

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYVASSWFA 10  
|||:|  
Db 215 PIHYCARWFA 224

## RESULT 7

ID Q920E8 PRELIMINARY; PRT; 120 AA.  
AC Q920E8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Perin-mimicking anti-idiotope heavy chain variable region  
DE (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Atkin J.D., Iape A., Jennings I.G., Horaitis O., Cotton R.G.H.;  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF307936; AAL09420.1; -.  
DR HSP; P01751; INQB.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
FT NON\_TER 1  
FT NON\_TER 120 120  
SQ SEQUENCE 120 AA; 13204 MW; DC4834AB1DB56F3C CRC64;

Query Match 61.5%; Score 40; DB 2; Length 120;  
Best Local Similarity 66.7%; Pred. No. 51;  
Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

Qy 2 IYVASS--WFA 11  
|||:|  
Db 99 IYGNPWFAY 110

## RESULT 8

ID Q8FCY4 PRELIMINARY; PRT; 153 AA.  
AC Q8FCY4;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein c4115.  
GN OrderedLocustNames=c4115;  
OS Escherichia coli O6.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=217992;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;  
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roessch P.,  
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
RT "Extensive mosaic structure revealed by the complete genome sequence  
of uropathogenic Escherichia coli.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).  
DR EMBL; AE016767; AA82553.1; -.  
KW Complete proteome.

Qy 133 PIYVASSWFA 11  
|||:|  
Db 139 FASTWFA 146

## RESULT 9

ID Q6MDY6 PRELIMINARY; PRT; 340 AA.  
AC Q6MDY6;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocustNames=pc0489;  
OS Parachlamydia sp. (strain UWR25) (subsp. Acanthamoeba sp.).  
OC Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae; Parachlamydia.  
OX NCBI\_TaxID=264201;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,  
RA Fartmann B., Brandt P., Nyakatura G.J., Droege M., Frishman D.,  
RA Rattei T., Mewes H.-W., Wagner M.;  
RT "Genome sequence of an amoeba symbiont and its use for reconstructing  
the evolutionary history of chlamydiae.";  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BX908798; CAF23213.1; -.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 340 AA; 39775 MW; 7313DCAEF62D10BE CRC64;

Query Match 61.5%; Score 40; DB 2; Length 340;  
Best Local Similarity 70.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IYVASSWFA 11  
|||:|  
Db 274 IYFSSWANY 283

## RESULT 10

ID Q7NV2 PRELIMINARY; PRT; 364 AA.  
AC Q7NV2;  
DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Similarities with unknown protein.  
GN OrderedLocustNames=plu0999;  
OS Photorhabdus luminescens (subsp. laumondii).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Photorhabdus.  
OX NCBI\_TaxID=141679;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TT01;  
RX MEDLINE=22957627; PubMed=14528314;  
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,  
RA Taouit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,  
RA Dassa E., Derose R., Dertelle S., Freysinet G., Gaudriault S.,  
RA Medigue C., Lanois A., Powell K., Sigulier P., Vincent R., Wingate V.,  
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;  
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus  
luminescens.";  
RL Nat. Biotechnol. 21:1307-1313 (2003).  
DR EMBL; BX571862; CAE13294.1; -.  
DR Photolista; plu0999; -.  
DR InterPro; IPR007113; Cupin region.  
DR InterPro; IPR011051; RmlC\_like\_cupin.  
KW Complete proteome.

Qy 61.5%; Score 40; DB 2; Length 364;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 5 ASSWFA 11
Db 254 ASSWFA 260

RESULT 11
Q9RJ56 Q9RJ56 PRELIMINARY; PRT; 374 AA.
AC Q9RJ56;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein SCO1652.
GN ORFNames=SC141.35c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939109; CAB59505.1; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000644; CBS.
DR InterPro; IPR008915; Peptidase_M50.
DR Pfam; PF00571; CBS; 2.
DR SMART; SM00116; Peptidase_M50; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 374 AA; 39795 MW; 24F9B18BC422C2F8 CRC64;

Query Match 61.5%; Score 40; DB 2; Length 374;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PIYVASSWF 9
Db 8 PVIVAPSWF 16

RESULT 12
Q828J6 Q828J6 PRELIMINARY; PRT; 376 AA.
AC Q828J6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV6673;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;

```

```

RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005047; BAC74384.1; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000644; CBS.
DR InterPro; IPR008915; Peptidase_M50.
DR Pfam; PF00571; CBS; 2.
DR SMART; SM00116; Peptidase_M50; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 376 AA; 40142 MW; A1429711C31A42EA CRC64;

Query Match 61.5%; Score 40; DB 2; Length 376;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PIYVASSWF 9
Db 8 PVIVAPSWF 16

RESULT 13
Q9JST2 Q9JST2 PRELIMINARY; PRT; 389 AA.
AC Q9JST2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein NMA2148.
GN OrderedLocusNames=NMA2148;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteriae; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506(2000).
DR EMBL; AL162758; CAB85360.1; -.
DR PIR; A81787; A81787.
DR InterPro; IPR008940; Prenyl_trans.
DR InterPro; IPR001440; TPR.
DR SMART; SM00028; TPR; 3.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 389 AA; 44448 MW; AEB0A5734EDAB65 CRC64;

```

Query Match 61.5%; Score 40; DB 2; Length 389;  
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 PIYYASSWFA 10  
 ||::|| |||  
 Db 16 PVFFAMGWFA 25

## RESULT 14

ID Q9K148 PRELIMINARY; PRT; 389 AA.  
 AC Q9K148; (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein NMB0339.  
 GN OrderedLocusNames=NMB0339;  
 OS Neisseria meningitidis (serogroup B).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 OC Neisseriaceae; Neisseria.  
 OX NCBI\_TaxID=491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MC58 / Serogroup B;  
 RX MEDLINE=20175755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;  
 RA Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,  
 RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,  
 RA Dodson R.J., Nelson W.C., Winn M.L., DeBoy R.T., Peterson J.D.,  
 RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,  
 RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,  
 RA Cittoni H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,  
 RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Maignani V., Pizza M.,  
 RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.";  
 RL Science 287:1809-1815(2000).  
 DR EMBL; AE002390; AAF40782.1; -.  
 DR PIR; F81211; F81211.  
 DR TTGR; NMB0339; -.  
 DR InterPro; IPR008940; Prenyl trans.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 389 AA; 44577 MW; DBA8F3BA22DEA517 CRC64;

Query Match 61.5%; Score 40; DB 2; Length 389;  
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 PIYYASSWFA 10  
 ||::|| |||  
 Db 16 PVFFAMGWFA 25

## RESULT 15

ID Q6LJ85 PRELIMINARY; PRT; 465 AA.  
 AC Q6LJ85;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Hypothetical protein.  
 GN OrderedLocusNames=PPRR0773;  
 OS Photobacterium profundum (Photobacterium sp. (strain S99)).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Photobacterium.  
 OX NCBI\_TaxID=74109;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,  
 RA Cestaro A., Malacrida G., Simionati B., Cannata N., Bartlett D.,  
 RA Valle G.;  
 RT "Genome analysis of Photobacterium profundum reveals the complexity of

high pressure adaptations.";  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -I- FUNCTION: Part of a binding-protein-dependent transport system.  
 CC Probably responsible for the translocation of the substrate across  
 CC the membrane (By similarity).  
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -I- SIMILARITY: Belongs to the binding-protein-dependent transport  
 CC system permease family.  
 DR EMBL; CR378677; CAG22645.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005215; P:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000515; BPD transp.  
 DR Pfam; PF00528; BPD transp.1; 1.  
 DR PROSITE; PS00928; ABC\_TM1; 1.  
 KW Complete proteome; Hypothetical protein; Transmembrane; Transport.  
 SQ SEQUENCE 465 AA; 52332 MW; E1F15F8646BE820 CRC64;

Query Match 61.5%; Score 40; DB 2; Length 465;  
 Best Local Similarity 66.7%; Pred. No. 2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PIYYASSWFP 9  
 |||| ||::||  
 Db 199 PIYIAAAWF 207

## RESULT 16

ID Q6BYK0 PRELIMINARY; PRT; 519 AA.  
 AC Q6BYK0;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Similar to CA57561|IPF1065 Candida albicans IPF1065 unknown  
 DE function.  
 GN ORFNames=DEHA0A09163g;  
 OS Debaryomyces hansenii CBS767.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.  
 OX NCBI\_TaxID=284592;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RG Genolevures;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantraye F., Hennequin C., Jaumiaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicaud J.M., Nikolski M., Ostas S., Ozier-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts.";  
 RL Nature 430:35-44(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS767;  
 RG Genoscope;  
 RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR382133; CAG84678.1; -.  
 DR InterPro; IPR008934; AcPase\_VanPerase.  
 DR InterPro; IPR010291; DUF895.  
 DR Pfam; PF05978; DUF895; 1.  
 SQ SEQUENCE 519 AA; 58022 MW; 2E5130975267758A CRC64;

Query Match 61.5%; Score 40; DB 2; Length 519;

```
Best Local Similarity 55.6%; Pred. No. 2.2e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIYVASSWF 9
Db 283 PMFASNWF 291
|:::|::|

RESULT 17
Q747F4 PRELIMINARY; PRT; 535 AA.
AC Q747F4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Thiolase, putative.
GN OrderedLocusNames=GSU3313;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.P., Wu D., Wu M., Ward N.L., Beanan M.J., Dodson R.J.,
RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R.,
RA Van Aken S.E., Lovley D.R., Fraser C.M.;
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT environments.";
RL Science 302:1967-1969 (2003).
DR EMBL; AE017180; AAR36703.1; -.
DR TIGR; GSU3313; -.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00108; Thiolase_N; 1.
DR Complete proteome.
SQ SEQUENCE 535 AA; 58200 MW; 81A09F9F27571007 CRC64;

Query Match 61.5%; Score 40; DB 2; Length 535;
Best Local Similarity 60.08; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYVASSWFA 10
Db 10 PVTVAASWMA 19
|:::|::|

RESULT 18
Q9GQRO PRELIMINARY; PRT; 2174 AA.
AC Q9GQRO;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extracellular matrix protein papilin precursor.
GN Name=Ppn;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=dp cn bw;
RX MEDLINE=20530499; PubMed=11076767;
RA Kramerova I.A., Kawaguchi N., Nelson R.E., Fessler L.I., Chen Y.,
RA Kramerov A.A., Kusche-Gullberg M., Kramer J.M., Ackley B.D.,
RA Stierov A.L., Prockop D.J., Fessler J.H.;
RT "Papilin in development; a pericellular protein with a homology to the
```

```
RT ADAMTS metalloproteinases";
RL Development 127:5475-5485 (2000).
CC -1- SIMILARITY: Contains 3 BPTI/Kunitz inhibitor domains.
DR EMBL; AF205357; AAG37995.1; -.
DR HSSP; P12111; 1KTH.
DR FlyBase; FBgn0003137; Ppn.
DR GO; GO:0005604; C:basement membrane; IDA.
DR InterPro; IPR010294; ADAM_spacer1.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR010909; PLAC.
DR InterPro; IPR002223; Prot_inh_Kunz-m.
DR InterPro; IPR000884; TSP1_inh_Kunz-m.
DR InterPro; IPR008197; WAP.
DR Pfam; PF05986; ADAM_spacer1; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00014; Kunitz_BPTI; 3.
DR Pfam; PF00090; TSP_1; 5.
DR Pfam; PF00095; WAP_1.
DR PRINTS; PR00003; 4DISULPHCORE.
DR PRINTS; PR00759; BASICPTASE.
DR SMART; SM00131; KU; 3.
DR SMART; SM00209; TSP1; 7.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4-DISULFIDE CORE; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 3.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 3.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00835; IG LIKE; 2.
DR PROSITE; PS00900; PLAC; 1.
DR PROSITE; PS00092; TSP1; 5.
DR PROSITE; PS00092; TSP1; 5.
DR Matrix protein; Signal.
FT SIGNAL 1 26 Potential.
SQ SEQUENCE 2174 AA; 231935 MW; 038F707952623120 CRC64;

Query Match 61.5%; Score 40; DB 2; Length 2174;
Best Local Similarity 45.5%; Pred. No. 9.6e+02;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYVASSWFAY 11
Db 276 PMFANSWNY 286
|:::|::|

RESULT 19
Q9VAV4 PRELIMINARY; PRT; 2772 AA.
AC Q9VAV4; Q9VAV3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG33103-PB.
GN Name=Ppn; ORFNames=CG33103;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley B.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
```

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lau X., Mattei B., McIntosh A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Matthei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy C., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheier F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RN Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E.E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
RT melanogaster euchromatic genome sequence.";  
RN Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
RT a genomics perspective.";  
RN Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review.";  
RN Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Contains 10 BPTI/Kunitz inhibitor domains.  
DR EMBL; AE003765; AAF56794.3; -;  
DR HSSP; P12111; 1KTH.

DR FlyBase; FBgn0003137; Ppn.  
DR GO; GO:0005604; C:basement membrane; IDA.  
DR InterPro; IPR010294; ADAM spacer1.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR010909; PLAC.  
DR InterPro; IPR002223; Prot\_Inh\_Kunz-m.  
DR InterPro; IPR008884; TSP1\_Inh\_Kunz-m.  
DR InterPro; IPR008197; WAP.  
DR Pfam; PF05986; ADAM\_spacer1; 1.  
DR Pfam; PF00047; Ig; 2.  
DR Pfam; PF00014; Kunitz\_BPTI; 10.  
DR Pfam; PF00090; TSP 1; 5.  
DR Pfam; PF00095; WAP; 1.  
DR PRINTS; PR00003; 4DISULPHCORE.  
DR PRINTS; PR00759; BASICPTASE.  
DR ProDom; PD000222; Prot\_Inh\_Kunz-m; 10.  
DR SMART; SM00408; IGC2; 3.  
DR SMART; SM00131; KU; 10.  
DR SMART; SM00209; TSP1; 7.  
DR SMART; SM00217; WAP; 1.  
DR PROSITE; PS00317; 4DISULFIDE\_CORE; 1.  
DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 9.  
DR PROSITE; PS00279; BPTI\_KUNITZ\_2; 10.  
DR PROSITE; PS00022; EGF 1; UNKNOWN\_1.  
DR PROSITE; PS00835; IG LIKE; 3.  
DR PROSITE; PS00900; PLAC; 1.  
DR PROSITE; PS00092; TSP1; 5.  
SQ SEQUENCE 2772 AA; 299154 MW; 3965DC92D30CCAAA CRC64;  
  
Qy 1 PIYYASSWFAY 11  
Db 276 PMPFANSWMY 286  
  
RESULT 20  
Q869A0 PRELIMINARY; PRT; 2776 AA.  
AC Q869A0;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Extracellular matrix protein papilin 2.  
GN NamesPpn;  
OS *Drosophila melanogaster* (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; *Drosophila*.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22552133; PubMed=12666201; DOI=10.1002/dvdy.10265;  
RA Kramerova I.A., Kramerov A.A., Fessler J.H.;  
RT "Alternative splicing of papilin and the diversity of *Drosophila*  
RT extracellular matrix during embryonic morphogenesis.";  
RL Dev. Dyn. 226:634-642(2003).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kramerova I., Fessler J.H.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Contains 10 BPTI/Kunitz inhibitor domains.  
DR EMBL; AF529179; AAC084907.1; -;  
DR HSSP; P12111; 1KTH.  
DR FlyBase; FBgn0003137; Ppn.  
DR GO; GO:0005604; C:basement membrane; IDA.  
DR InterPro; IPR010294; ADAM spacer1.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR007110; Ig-like.

Query Match 61.5%; Score 40; DB 2; Length 2772;  
Best Local Similarity 45.5%; Pred. No. 1.2e+03;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;



DR InterPro: IPR003598; Ig\_c2.  
DR InterPro: IPR010909; PLAC.  
DR InterPro: IPR02223; Prot\_Inh\_Kunz-m.  
DR InterPro: IPR00884; TSP1.  
DR InterPro: IPR008197; WAP.  
DR Pfam: PF05986; ADAM\_spacer1; 1.  
DR Pfam: PF00047; Ig\_2.  
DR Pfam: PF00014; Kunitz\_BPTI; 10.  
DR Pfam: PF00090; TSP\_1; 5.  
DR Pfam: PF00095; WAP; 1.  
DR PRINTS: PR00003; 4DISULPHCORE.  
DR PRINTS: PR00759; BASICPTASE.  
DR ProDom: PD000222; Prot\_Inh\_Kunz-m; 10.  
DR SMART: SM00408; IGC2; 3.  
DR SMART: SM00131; KU; 10.  
DR SMART: SM00209; TSP1; 7.  
DR SMART: SM00217; WAP; 1.  
DR PROSITE: PS00317; 4\_DISULFIDE\_CORE; 1.  
DR PROSITE: PS00280; BPTI\_KUNITZ\_1; 9.  
DR PROSITE: PS02779; BPTI\_KUNITZ\_2; 10.  
DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE: PS50835; IG\_LIKE; 3.  
DR PROSITE: PS50900; PLAC; 1.  
DR PROSITE: PS50092; TSP1; 5.  
KW Matrix protein.  
SQ SEQUENCE 2776 AA; 92D38A17360D2D42 CRC64;  
Query Match 61.5%; Score 40; DB 2; Length 2776;  
Best Local Similarity 45.5%; Pred. No. 1.2e+03;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
QY 1 PIYASSWPAY 11  
DB 276 PMFFANSWNY 286  
RESULT 21  
Q7KRX2 PRELIMINARY; PRT; 2894 AA.  
ID Q7KRX2  
AC Q7KRX2;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE CG33103-PA.  
GN Name=Fpn; ORFNames=CG33103;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=10731132; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Branton R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective.";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Mira S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a systematic review.";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Contains 12 BPTI/Kunitz inhibitor domains.  
DR EMBL; AE003765; AAF56795.3; -.  
DR HSSP; P10646; 1ADZ.  
DR GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.  
DR InterPro: IPR010294; ADAM\_spacer1.  
DR InterPro: IPR006209; EGF\_like.  
DR InterPro: IPR003599; IG.  
DR InterPro: IPR007110; IG-like.  
DR InterPro: IPR003598; IG\_c2.  
DR InterPro: IPR010909; PLAC.  
DR InterPro: IPR002223; Prot\_Inh\_Kunz-m.  
DR InterPro: IPR000884; TSP1.  
DR InterPro: IPR008197; WAP.

```
DR Pfam; PF05986; ADAM_spacer1; 1.
DR Pfam; PF00047; Ig; 2.
DR Pfam; PF00014; Kunitz_BPTI; 12.
DR Pfam; PF00090; TSP_1; 5.
DR Pfam; PF00095; WAP; 1.
DR PRINTS; PR00003; 4DISULPHCORE.
DR PRINTS; PR00759; BASICPTASE.
DR PRODOM; PD000222; Prot_Inh_Kunz-m; 12.
DR SMART; SM00408; IGC2; 3.
DR SMART; SM00131; KU; 12.
DR SMART; SM00209; TSP1; 7.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4_DISULFIDE_CORE; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 11.
DR PROSITE; PS0279; BPTI_KUNITZ_2; 12.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS0835; IG_LIKE; 3.
DR PROSITE; PS0900; PLAC; 1.
DR PROSITE; PS0092; TSP1; 5.
DR PROSITE; PS0092; TSP1; 5.
KW Matrix Protein.
SQ SEQUENCE 2894 AA; 312663 MW; 1BFBFEBAD9B214BC CRC64;

Query Match 61.5%; Score 40; DB 2; Length 2894;
Best Local Similarity 45.5%; Pred. No. 1.3e+03;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASSWFAY 11
|:::|::|
Db 276 PMFFANSWNWY 286

RESULT 22
Q86829 PRELIMINARY; PRT; 2898 AA.
ID O86829;
AC O86829;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Extracellular matrix protein papilin 3.
GN Name=Ppn;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22552133; PubMed=12666201; DOI=10.1002/dvdy.10265;
RA Kramerova I.A., Kramerov A.A., Fessler J.H.;
RT "Alternative splicing of papilin and the diversity of Drosophila
RT extracellular matrix during embryonic morphogenesis.";
RL Dev. Dyn. 226:634-642(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Kramerova I., Fessler J.H.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 12 BPTI/Kunitz inhibitor domains.
EMBL; AF529180; AAC84908.1; -.
DR HSP; P12111; 1KTH.
DR FlyBase; FBgn0003137; Ppn.
DR GO; GO:0005604; C:basement membrane; IDA.
DR InterPro; IPR010294; ADAM_spacer1.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR010909; PLAC.
DR InterPro; IPR002223; Prot_Inh_Kunz-m.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008197; WAP.
DR Pfam; PF05986; ADAM_spacer1; 1.
DR Pfam; PF00047; Ig; 2.
DR Pfam; PF00014; Kunitz_BPTI; 12.
DR Pfam; PF00090; TSP_1; 5.
```

```
DR Pfam; PF00095; WAP; 1.
DR PRINTS; PR00003; 4DISULPHCORE.
DR PRINTS; PR00759; BASICPTASE.
DR PRODOM; PD000222; Prot_Inh_Kunz-m; 12.
DR SMART; SM00408; IGC2; 3.
DR SMART; SM00131; KU; 12.
DR SMART; SM00209; TSP1; 7.
DR SMART; SM00217; WAP; 1.
DR PROSITE; PS00317; 4_DISULFIDE_CORE; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 11.
DR PROSITE; PS0279; BPTI_KUNITZ_2; 12.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS0835; IG_LIKE; 3.
DR PROSITE; PS0900; PLAC; 1.
DR PROSITE; PS0092; TSP1; 5.
KW Matrix Protein.
SQ SEQUENCE 2898 AA; 313250 MW; 2F992742F2D64A00 CRC64;

Query Match 61.5%; Score 40; DB 2; Length 2898;
Best Local Similarity 45.5%; Pred. No. 1.3e+03;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASSWFAY 11
|:::|::|
Db 276 PMFFANSWNWY 286

RESULT 23
LYST_MOUSE
ID _LYST_MOUSE STANDARD; PRT; 3788 AA.
AC P97412; Q62403; Q8VBS6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lysosomal trafficking regulator (Beige protein) (CHS1 homolog).
GN Name=Chsi; Synonyms=bg, Lyst;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=97358584; PubMed=9215680; DOI=10.1093/hmg/6.7.1091;
RA Barbosa M.D.F.S., Barrat F.J., Tchernev V.T., Nguyen J.A.,
RA Mishra V.S., Colman S.D., Pastural E., Dufourcq-Lagelouse R.,
RA Fischer A., Holcombe R.F., Wallace M.R., Brandt S.J.,
RA De Saint Basile G., Kingsmore S.F.;
RT "Identification of mutations in two major mRNA isoforms of the
RT Chediak-Higashi syndrome gene in human and mouse.";
RL Hum. Mol. Genet. 6:1091-1098(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX STRAIN=C57BL/6J;
RX MEDLINE=96353977; PubMed=8717042; DOI=10.1038/382262a0;
RA Barbosa M.D.F.S., Nguyen Q.A., Tchernev V.T., Ashley J.A.,
RA Dettler J.C., Blaydes S.M., Brandt S.J., Chotai D., Hodgman C.,
RA Solari R.C.E.S., Lovett M., Kingsmore S.F.;
RT "Identification of the homologous beige and Chediak-Higashi syndrome
RT genes.";
RL Nature 382:262-265 (1996).
RN [3]
RP ERRATUM.
RA Barbosa M.D.F.S., Nguyen Q.A., Tchernev V.T., Ashley J.A.,
RA Dettler J.C., Blaydes S.M., Brandt S.J., Chotai D., Hodgman C.,
RA Solari R.C.E.S., Lovett M., Kingsmore S.F.;
RL Nature 385:97-97(1997).
RN [4]
RP SEQUENCE OF 1428-3788 FROM N.A. (ISOFORM 1).
RX STRAIN=C57BL/6J;
RX MEDLINE=96259558; PubMed=8673129;
RA Perou C.M., Moore K.J., Nagle D.L., Misumi D.J., Woolf E.A.,
RA McGrail S.H., Holmgren L., Brody T.B., Dussault B.J., Monroe C.A.,
RA Duyk G.M., Pryor R.J., Li L., Justice M.J., Kaplan J.;
```

```
RT "Identification of the murine beige gene by YAC complementation and
RL positional cloning."
CC Nat. Genet. 13:303-308(1996).
CC -!- FUNCTION: May be required for sorting endosomal resident proteins
CC into late multivesicular endosomes by a mechanism involving
CC microtubules.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=P97412-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P97412-2; Sequence=VSP 006783, VSP 006784;
CC -!- DISEASE: Defects in Chai are the cause of Beige, an autosomal
CC recessive disorder characterized by hypopigmentation, bleeding,
CC immune cell dysfunction, abnormal intracellular transport to and
CC from the lysosome, and giant inclusion bodies in a variety of cell
CC types.
CC -!- SIMILARITY: Contains 1 BEACH domain.
CC -!- SIMILARITY: Contains 7 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: U70015; AAC53011.1; -
EMBL: L77884; AAL40134.1; -
EMBL: U52461; AAB60778.1; -
DR PIR: T30851; T30851.
DR HSSP: Q8NFP9; 1M11.
DR MGI: MGI:107448; Lyst.
DR InterPro: IPR000409; Beige BEACH.
DR InterPro: IPR001680; WD40__like.
DR InterPro: IPR011046; WD40__like.
DR Pfam: PF02138; Beach; 1.
DR Pfam: PF00400; WD40; 5.
DR SMART: SM00320; WD40; 4.
DR PROSITE: PS0197; BEACH; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 1.
DR PROSITE: PS00678; WD_REPEATS_2; 1.
DR PROSITE: PS0082; WD_REPEATS_2; 1.
DR PROSITE: PS0294; WD_REPEATS_REGION; 1.
DR KW Alternative splicing; Protein transport; Repeat; Transport; WD repeat.
FT REPEAT 662 700 WD 1.
FT REPEAT 1576 1620 WD 2.
FT DOMAIN 3126 3409 BEACH.
FT REPEAT 3350 3589 WD 3.
FT REPEAT 3601 3640 WD 4.
FT REPEAT 3601 3640 WD 4.
FT REPEAT 3643 3686 WD 5.
FT REPEAT 3687 3731 WD 6.
FT REPEAT 3736 3775 WD 7.
FT VARSPLIC 1509 1545
FT EGDPREVESINPGRLIEDGCIHLISLGSKALMIQV ->
FT GMAAGSDYTKILQIACLSFPHIWDQVNFVKCYSP (in
FT isoform 2).
FT /FTId=VSP 006783.
FT Missing (in isoform 2).
FT /FTId=VSP 006784.
FT VARSPLIC 1546 3788
FT SEQUENCE 3788 AA; 425283 MW; F01BE837C676A750 CRC64;
Query Match 61.5%; Score 40; DB 1; Length 3788;
Best Local Similarity 75.0%; Pred. No. 1.7e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PIYVASSW 8
Db 2927 PIYPTSM 2934
RESULT 24
```

```
Q9Z2X9
ID Q9Z2X9 PRELIMINARY; PRT; 3788 AA.
AC Q9Z2X9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BEIGE
GN Name=beige;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA; TISSUE=Spleen;
RX MEDLINE=99315327; PubMed=10384041;
RA Mori N., Nishikawa T., Higuchi K., Nishimura M.;
RT "Deletion in the beige gene of the beige rat owing to recombination
RL between LINEs.";
RL Mamm. Genome 10:692-695(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DA; TISSUE=Spleen;
RA Mori M., Nishikawa T., Higuchi K., Nishimura M.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 5 WD repeats.
DR EMBL: AB020019; BAA34688.1; -.
DR PIR: T13960; T13960.
DR HSSP: Q8NFP9; 1M11.
DR InterPro: IPR000409; Beige BEACH.
DR InterPro: IPR001680; WD40__like.
DR InterPro: IPR011046; WD40__like.
DR Pfam: PF02138; Beach; 1.
DR Pfam: PF00400; WD40; 5.
DR SMART: SM00320; WD40; 4.
DR PROSITE: PS0197; BEACH; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
DR PROSITE: PS0082; WD_REPEATS_2; 1.
DR PROSITE: PS0294; WD_REPEATS_REGION; 1.
DR KW Repeat; WD repeat.
SQ SEQUENCE 3788 AA; 426601 MW; E51FB90F0E2E6550 CRC64;
Query Match 61.5%; Score 40; DB 2; Length 3788;
Best Local Similarity 75.0%; Pred. No. 1.7e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PIYVASSW 8
Db 2927 PIYPTSM 2934
RESULT 25
Q97699
ID Q97699 PRELIMINARY; PRT; 3796 AA.
AC Q97699;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lysosomal trafficking regulator.
GN Name=LYST;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC Yamakuchi H., Hirano T., Hara K., Sugimoto Y.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Agaba M.K., Yamakuchi H., Takasuga A., Takeda H., Hirano T.,
RA Sugimoto Y.;
```

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 5 WD repeats.  
 DR EMBL; AF114785; AADI3618.1; -.  
 DR PIR; T18514; T18514.  
 DR HSP; Q8NFP9; IM1.  
 DR InterPro; IPR000409; Beige\_BEACH.  
 DR InterPro; IPR000923; BlueCu\_1.  
 DR InterPro; IPR001680; WD40.  
 DR InterPro; IPR011046; WD40\_like.  
 DR Pfam; PF021138; Beach; 1.  
 DR Pfam; PF00400; WD40; 5.  
 DR SMART; SM00320; WD40; 4.  
 DR PROSITE; PS00197; BEACH; 1.  
 DR PROSITE; PS00196; COPPER\_BLUE; UNKNOWN\_1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_1.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 1.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 3796 AA; 428539 MW; 10C9FE03E2DF086E CRC64;

Query Match 61.5%; Score 40; DB 2; Length 3796;  
 Best Local Similarity 75.0%; Pred. No. 1.7e+03;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASW 8  
 ||||| :||  
 Db 2935 PIYYPTSW 2942

RESULT 26  
 ID Q9TTK4 PRELIMINARY; PRT; 3796 AA.

AC Q9TTK4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Lysosomal trafficking regulator.  
 GN Name=LYST;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20063686; PubMed=10594238;  
 RA Kunieda T., Nakagiri M., Takami M., Ide H., Ogawa H.;  
 RT "Cloning of bovine LYST gene and identification of a missense mutation  
 associated with Chediak-Higashi syndrome of cattle.";  
 RL Mamm. Genome 10:1146-1149(1999).  
 CC -1- SIMILARITY: Contains 5 WD repeats.  
 DR EMBL; AB026620; BAA77267.1; -.  
 DR HSP; Q8NFP9; IM1.  
 DR InterPro; IPR000409; Beige\_BEACH.  
 DR InterPro; IPR000923; BlueCu\_1.  
 DR InterPro; IPR001680; WD40.  
 DR InterPro; IPR011046; WD40\_like.  
 DR Pfam; PF021138; Beach; 1.  
 DR Pfam; PF00400; WD40; 5.  
 DR SMART; SM00320; WD40; 4.  
 DR PROSITE; PS00197; BEACH; 1.  
 DR PROSITE; PS00196; COPPER\_BLUE; UNKNOWN\_1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_1.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 1.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 3796 AA; 428537 MW; 4AC00B9928E882B7 CRC64;

Query Match 61.5%; Score 40; DB 2; Length 3796;  
 Best Local Similarity 75.0%; Pred. No. 1.7e+03;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PIYYASW 8  
 ||||| :||  
 Db 2935 PIYYPTSW 2942

RESULT 27  
 ID Q99698 STANDARD; PRT; 3801 AA.

LYST\_HUMAN  
 AC Q99698; Q43274; Q96TD7; Q96TD8; Q99709; Q9H133;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Lysosomal trafficking regulator (Beige homolog).  
 GN Name=CHS1; Synonyms=CHS, LYST;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC MEDLINE=97051925; PubMed=8896560;  
 RA Nagle D.L., Karim M.A., Woolf E.A., Holmgren L., Bork P., Misumi D.J.,  
 RA McGrail S.H., Duseault B.J., Perou C.M., Boissy R.E., Duyk G.M.,  
 RA Spritz R.A., Moore K.J.;  
 RT "Identification and mutation analysis of the complete gene for  
 Chediak-Higashi syndrome.";  
 RL Nat. Genet. 14:307-311(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Liver;  
 RX MEDLINE=96353977; PubMed=8717042; DOI=10.1038/382262a0;  
 RA Barbosa M.D.F.S., Nguyen Q.A., Tchernev V.T., Ashley J.A.,  
 RA Dettler J.C., Blaydes S.M., Brandt S.J., Chotai D., Hodgman C.,  
 RA Solari R.C.E.S., Lovett M., Kingsmore S.F.;  
 RT "Identification of the homologous beige and Chediak-Higashi syndrome  
 genes.";  
 RL Nature 382:262-265 (1996).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RX MEDLINE=97358584; PubMed=9215680; DOI=10.1093/hmg/6.7.1091;  
 RA Barbosa M.D.F.S., Barrat F.J., Tchernev V.T., Nguyen Q.A.,  
 RA Mishra V.S., Colman S.D., Pastural E., Dufourcq-Lagelouse R.,  
 RA Fischer A., Holcombe R.F., Wallace M.R., Brandt S.J.,  
 RA de Saint Basile G., Kingsmore S.F.;  
 RT "Identification of mutations in two major mRNA isoforms of the  
 Chediak-Higashi syndrome gene in human and mouse.";  
 RL Hum. Mol. Genet. 6:1091-1098(1997).  
 RN [4]  
 RP SEQUENCE OF 1-1475 FROM N.A.  
 RA Barrat F.J., Barbosa M.D.F.S., Pastural E., Dufourcq-Lagelouse R.,  
 RA Kingsmore S.F., de Saint Basile G.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 1-3186 FROM N.A. (ISOFORMS 1; 2 AND 3).  
 RA Cobley V.;  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SLIPICE ISOFORM(S) THAT ARE POTENTIAL NMD TARGET(S).  
 RX PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;  
 RA Hillman R.T., Green R.E., Brenner S.E.;  
 RT "An unappreciated role for RNA surveillance.";  
 RL Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).  
 RN [7]  
 RP VARIANTS CHS HIS-1563 AND ASP-1999.  
 RX MEDLINE=21845150; PubMed=11857544; DOI=10.1002/ajmg.10184 abs;  
 RA Karim M.A., Suzuki K., Fukui K., Oh J., Nagle D.L., Moore K.J.,  
 RA Barbosa E., Falik-Borenstein T., Filipovich A., Ishida Y.,  
 RA Kivrikko S., Klein C., Kreuz F., Levin A., Miyajima H., Requeiro J.,  
 RA Russo C., Uyama E., Vierimaa O., Spritz R.A.;  
 RT "Apparent genotype-phenotype correlation in childhood, adolescent, and  
 adult Chediak-Higashi syndrome.";  
 RL Am. J. Med. Genet. 108:16-22(2002).  
 CC -1- FUNCTION: May be required for sorting endosomal resident proteins

into late multivesicular endosomes by a mechanism involving microtubules.

-!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).

-!- ALTERNATIVE PRODUCTS:

Event-Alternative splicing; Named isoforms=3;  
Comment=Additional isoforms seem to exist;  
Name=1;

isoId=099698-1; Sequence=Displayed;  
Note=May be produced at very low levels due to a premature stop codon in the mRNA, leading to nonsense-mediated mRNA decay;  
Name=2;

isoId=099698-2; Sequence=VSP\_006781, VSP\_006782;  
Name=3;

isoId=099698-3; Sequence=VSP\_006779, VSP\_006780;

-!- TISSUE SPECIFICITY: Abundantly expressed in adult and fetal thymus, peripheral blood leukocytes, bone marrow and several regions of the adult brain.

-!- DISEASE: Defects in CHS1 are the cause of Chediak-Higashi syndrome (CHS) [MIM:214500]. CHS is a rare autosomal recessive disorder characterized by hypopigmentation, severe immunologic deficiency, a bleeding tendency, neurologic abnormalities, abnormal intracellular transport to and from the lysosome, and giant inclusion bodies in a variety of cell types.

-!- SIMILARITY: Contains 1 BEACH domain.

-!- SIMILARITY: Contains 7 WD repeats.

-!- DATABASE: NAME=Mutations of the CHS1 gene;  
NOTE=Retina International's Scientific Newsletter;  
WWW="http://www.retina-international.com/sci-news/chsmut.htm".

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; U84744; ABA87737.1; -  
EMBL; U67615; ABA41309.1; -  
EMBL; U72192; AAB39697.1; -  
EMBL; L77889; AAB51608.1; -  
EMBL; U70064; ABA41533.1; -  
EMBL; AL121997; CAC17580.1; -  
EMBL; AL121997; CAC17581.1; -  
EMBL; AL121997; CAC17582.1; -  
HSP; Q8NPP9; IM11.  
Gene; HGNC:1968; CHS1.  
MIM; 214500; -  
GO; GO:0005768; C:endosome; TAS.  
GO; GO:0015630; C:microtubule cytoskeleton; TAS.  
GO; GO:0006968; P:cellular defense response; TAS.  
GO; GO:0008333; P:endosome to lysosome transport; TAS.  
GO; GO:0007165; P:signal transduction; TAS.  
InterPro; IPR000409; Beige\_BEACH.  
InterPro; IPR011036; PH-related.  
InterPro; IPR001680; WD40.  
InterPro; IPR011046; WD40\_like.  
Pfam; PF02138; Beach; 1.  
Pfam; PF00400; WD40; 5.  
ProDom; PD000018; WD40; 1.  
SMART; SM00320; WD40; 4.  
PROSITE; PS50197; BEACH; 1.  
PROSITE; PS00678; WD\_REPEATS\_1; 1.  
PROSITE; PS50082; WD\_REPEATS\_2; 1.  
PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
Alternative splicing; Disease mutation; Protein transport; Repeat; Transport; WD repeat.  
DOMAIN 31 36 Poly-Glu.  
DOMAIN 2448 2452 Poly-Leu.  
REPEAT 662 700 WD 1.  
REPEAT 1582 1626 WD 2.  
DOMAIN 3120 3422 BEACH.

FT REPEAT 3563 3602 WD 3.  
FT REPEAT 3614 3653 WD 4.  
FT REPEAT 3656 3699 WD 5.  
FT REPEAT 3700 3744 WD 6.  
FT REPEAT 3749 3788 WD 7.  
FT VARSPLIC 1515 1531  
ESDRPEGAETINPGERL -> GMMTGLSLTYTKIVFRL  
(in isoform 3).  
/FTid=VSP\_006779.  
Missing (in isoform 3).  
/FTid=VSP\_006780.  
VCRSPFKIIAEVLG -> MARSFRKRCQSCT (in  
isoform 2).  
/FTid=VSP\_006781.  
Missing (in isoform 2).  
/FTid=VSP\_006782.  
R -> H (in CHS).  
/FTid=VAR\_013556.  
V -> D (in CHS).  
/FTid=VAR\_013557.  
SQ SEQUENCE 3801 AA; 429117 MW; 288B370AD4BDAF11 CRC64;  
Query Match 61.5%; Score 40; DB 1; Length 3801;  
Best Local Similarity 75.0%; Pred. No. 1.7e+03;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 PIYYASSW 8  
||||| :||  
Db 2940 PIYYPTSW 2947

RESULT 28  
Q6C293 PRELIMINARY; PRT; 879 AA.  
AC Q6C293  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Yarrowia lipolytica chromosome F of strain CLIB99 of Yarrowia  
DE lipolytica.  
GN ORFNames=YALI0F09691g;  
OS Yarrowia lipolytica CLIB99.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Dipodascaceae; Yarrowia.  
OX NCBI\_TaxID=284591;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CLIB99;  
RG Genolevures;  
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talia E.,  
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
RA Bolrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,  
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
RA Kerrest A., Koszul K., Lemaire M., Lesur I., Ma L., Muller H.,  
RA Nicaud J.M., Nikoleki M., Oztas S., Ozier-Kalogeropoulos O.,  
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
RA Swennene D., Tekala F., Wesolowski-Louvet M., Westhof E., Wirth B.,  
RA Zenou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,  
RA Wincker P., Souciet J.L.;  
RT "Genome evolution in yeasts."  
RL Nature 430:35-44(2004).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CLIB99;  
RG Genoscope;  
RA EMBL; CR382132; CAG78026.1; -  
DR EMBL; CR382132; CAG78026.1; -  
DR InterPro; IPR004648; Oligopept\_transpt.  
DR InterPro; IPR004813; Tetrpept\_transpt.  
DR Pfam; PF03169; OPT; 1.

```

DR TIGRFAMs; TIGR00727; ISP4_OPT; 1.
DR TIGRFAMs; TIGR00728; OPT_SFam; 1.
SQ SEQUENCE 879 AA; 101056 MW; 0FF62B08C4A41510 CRC64;

Query Match 60.8%; Score 39.5; DB 2; Length 879;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

Qy 1 PIYASSW--FAY 11
|||||
Db 257 PIYQNSWAGGFGY 270

RESULT 29
Q6H422 PRELIMINARY; PRT; 190 AA.
AC Q6H422
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Putative fiber protein FB34.
GN Name=P0651G05.38; Synonyms=P0645D04.18;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, PAC
RT clone:P0651G05."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, PAC
RT clone:P0645D04."
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006528; BAD26527.1; -.
DR EMBL; AP006727; BAD29644.1; -.
DR InterPro; IPR009606; DUF1218.
DR Pfam; PF06749; DUF1218; 1.
SQ SEQUENCE 190 AA; 20812 MW; D50F600B3314C095 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 190;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IYASSWFAY 11
|||||
Db 97 IYFASSWVTF 106

RESULT 30
PANE BACSU STANDARD; PRT; 298 AA.
AC O34661;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 23-OCT-2004 (Rel. 45, Last annotation update)
DE Probable 2-dehydropantoate 2-reductase (EC 1.1.1.169) (Ketopantoate
DE reductase) (KPA reductase).
GN Name=pane; Synonyms=apba; OrderedLocusNames=BSU15110;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA Bartero M., Presecan E., Glaser P., Richou A., Danchin A.;
RT "Bacillus subtilis chromosomal region downstream nprE."
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

```

---

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Avevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriest L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Etian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rose M., Reynolds S., Rieger M., Rivoita C., Rocha E., Roche B.,
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitenegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256(1997).
CC -I- FUNCTION: Catalyzes the NADPH-dependent reduction of ketopantoate
CC into pantoic acid (By similarity).
CC -I- CATALYTIC ACTIVITY: (R)-pantoate + NADP(+) = 2-dehydropantoate +
CC NADPH.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -I- SIMILARITY: Belongs to the ketopantoate reductase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z98682; CAB11364.1; -.
CC EMBL; Z99111; CAB13384.1; -.
CC PIR; F69875; F69875.
CC Subtilist; BGI3369; pane.
CC InterPro; IPR008927; 6GDH_C like.
CC InterPro; IPR003710; Apba.
CC Pfam; PF02558; Apba; 1.
CC TIGRFAMs; TIGR00745; apba_pane; 1.
CC Complete proteome; NADP; Oxidoreductase; Pantothenate biosynthesis.
FT NP BIND 7 12 NADP (potential).
FT ACT SITE 179 179 By similarity.
FT ACT SITE 258 258 By similarity.
SQ SEQUENCE 298 AA; 33287 MW; A0B645A0A415FFD5 CRC64;

Query Match 60.0%; Score 39; DB 1; Length 298;
Best Local Similarity 55.6%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYASSWF 9
|||||
Db 165 PIYETDWWY 173

```

```
RESULT 31
Q9M9D8      PRELIMINARY;      PRT;      399 AA.
AC Q9M9D8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE T16N11.7 protein.
GN Name=T16N11.7;
OS Arabidopsis thaliana (Mouse-ear cress);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federpiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Alrafi H., Araujo R., Huijar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC013453; AAF71981.1; --
DR PIR: E86289; E86289.
DR InterPro: IPR004252; Transposase_24.
DR Pfam: PF03004; Transposase_24; 1.
SQ SEQUENCE 399 AA; 44946 MW; C82416EFBE118B54 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 399;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PIYYASSWFA 10
|||.:||:
Db 98 PIFYKEAWFS 107

RESULT 32
Q6GR44      PRELIMINARY;      PRT;      400 AA.
AC Q6GR44;
DT 05-JUN-2004 (TrEMBLrel. 27, Created)
DT 05-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC81044 protein.
GN Name=MGC81044;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Beasak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
```

```
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE=Ovary;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE=Ovary;
RA Klein S., Gerhard D.S.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC071089; AAH71089.1; --
DR GO: GO:0004553; F:carbohydrate activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPR002241; Glyco_hydro_27.
DR InterPro: IPR000111; Glyco_hydro_GHD.
DR Pfam: PF02065; Melibiase; 1.
DR PRINTS: PR00740; GLHYDLASE27.
DR ProDom: PD02572; Glyco_hydro_GHD; 1.
DR PROSITE: PS00512; ALPHA_GALACTOSIDASE; 1.
SQ SEQUENCE 400 AA; 44854 MW; CBC4770F613FF3BA CRC64;

Query Match 60.0%; Score 39; DB 2; Length 400;
Best Local Similarity 63.6%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYYASSWFA 11
|||.:||:
Db 172 PLYSCSWPAY 182

RESULT 33
Q90744      PRELIMINARY;      PRT;      405 AA.
AC Q90744;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alpha-N-acetylgalactosaminidase (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Liver;
RX MEDLINE=34060104; PubMed=8241271; DOI=10.1016/0167-4781(93)90158-A;
RA Davis M.O., Hata J., Smith D., Walker J.C.;
RT "Cloning and sequence of a chicken alpha-N-acetylgalactosaminidase
RT gene.";
RL Biochim. Biophys. Acta 1216:296-298(1993).
DR EMBL: L18754; AAA16614.1; --
DR PIR: S45522; S45522.
DR PDB: 1KTB; X-ray; A=1-404.
DR PDB: 1KTC; X-ray; A=1-404.
DR GO: GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR Pfam: PF02065; Melibiase; 1.
DR PRINTS: PR00740; GLHYDLASE27.
DR ProDom: PD02572; Glyco_hydro_GHD; 1.
DR PROSITE: PS00512; ALPHA_GALACTOSIDASE; 1.
FT NON TER 1
SQ SEQUENCE 405 AA; 45615 MW; ELECO061739C305C CRC64;

Query Match 60.0%; Score 39; DB 2; Length 405;
Best Local Similarity 63.6%; Pred. No. 2.6e+02;
```

```
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYVASSWFAY 11
Db 166 PIVYSCSWPAY 176

RESULT 34
Q7Q6H3 Q7Q6H3 PRELIMINARY; PRT; 428 AA.
AC Q7Q6H3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP6545.
GN Name=agCG48798; ORFNames=ENSANGG0000014894;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008960; EAA11949.1; -.
DR HSP; P06280; I146.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR002241; Glyco_hydro_27.
DR InterPro; IPR000111; Glyco_hydro_GHD.
DR Pfam; PF02065; Melibiase; 1.
DR PRINTS; PR00740; GLHYDLASE27.
DR ProDom; PD002572; Glyco_hydro_GHD; 1.
DR PROSITE; PS00512; ALPHA_GALACTOSIDASE; 1.
SQ SEQUENCE 428 AA; 48797 MW; 113043C056D71320 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 428;
Best Local Similarity 63.6%; Pred. No. 2.7e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYVASSWFAY 11
Db 189 PIVYSCSWPAY 199

RESULT 35
Q6GQ17 Q6GQ17 PRELIMINARY; PRT; 431 AA.
AC Q6GQ17;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE LOC443592 protein (Fragment).
GN Name=LOC443592;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Dlatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
```

```
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC072931; AAH72931.1; -.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR002241; Glyco_hydro_27.
DR InterPro; IPR000111; Glyco_hydro_GHD.
DR Pfam; PF02065; Melibiase; 1.
DR PRINTS; PR00740; GLHYDLASE27.
DR ProDom; PD002572; Glyco_hydro_GHD; 1.
DR PROSITE; PS00512; ALPHA_GALACTOSIDASE; 1.
FT NON TER 1
SQ SEQUENCE 431 AA; 48787 MW; 4C578DB433983471 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 431;
Best Local Similarity 63.6%; Pred. No. 2.7e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PIYVASSWFAY 11
Db 205 PIVYSCSWPAY 215

RESULT 36
Q21801 Q21801 PRELIMINARY; PRT; 451 AA.
AC Q21801;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein R07B7.11.
GN ORFNames=R07B7.11;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99089613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RL investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
```



```
RA Harris B.R.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z75955; CAB00120.1; -.
DR PIR; T24018; 1KTB.
DR HSPP; Q90744; 1KTB.
DR WormBase; WBGene00011095; R07B7.11.
DR WormPep; R07B7.11; CE06273.
DR GO; GO:0004553; P:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR002241; Glyco_hydro_27.
DR InterPro; IPR000111; Glyco_hydro_GHD.
DR Pfam; PF02065; Melibiase; I.
DR PRINTS; PR00740; GLYDRLASE27.
DR PROSITE; PS00512; ALPHA_GALACTOSIDASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 451 AA; 51423 MW; F7BB2B05D5C6419 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 451;
Best Local Similarity 63.6%; Pred. No. 2.9e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIYVASSWPAY 11
|||: ||| |||
DB 180 PINTVCSWPAY 190

RESULT 37
ID O02259 PRELIMINARY; PRT; 500 AA.
AC O02259;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F44F1.1.
GN ORFNames=F44F1.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Dobson R.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z81083; CAB03099.2; -.
DR PIR; T22175; T22175.
DR WormBase; WBGene00009693; F44F1.1.
DR WormPep; F44F1.1; CE36159.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004198; P:calpain activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001300; Peptidase C2.
DR Pfam; PF00648; Peptidase C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00230; Cyapc; 1.
DR PROSITE; PS50203; CALPAIN_CAT; 1.
KW Hypothetical protein.
SQ SEQUENCE 500 AA; 57866 MW; A1A18E48DD9CB535 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 500;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 YASSWPAY 11
|||: ||| ||| |||
```

```
Db 219 YRSNWPAY 226

RESULT 38
Q98M81 PRELIMINARY; PRT; 520 AA.
ID Q98M81
AC Q98M81;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Trehalose-6-phosphate phosphatase.
GN OrderedLocusNames=ml10691;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303059;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002995; BAB48232.1; -.
DR HSPP; P31677; IGZ5.
DR GO; GO:0003825; F:alpha,alpha-trehalose-phosphate synthase (U. . . ; IEA.
DR GO; GO:0005992; P:trehalose biosynthesis; IEA.
DR InterPro; IPR001830; Glyco_transf_20.
DR Pfam; PF00982; Glyco_transf_20; 1.
KW Complete proteome.
SQ SEQUENCE 520 AA; 58996 MW; CE466A103191A963.CRC64;

Query Match 60.0%; Score 39; DB 2; Length 520;
Best Local Similarity 63.6%; Pred. No. 3.3e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIYVASSWPAY 11
|||: ||| |||
DB 126 PIFRSDWEAY 136

RESULT 39
Q96ZH6 PRELIMINARY; PRT; 530 AA.
ID Q96ZH6
AC Q96ZH6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein ST1858.
GN OrderedLocusNames=ST1858;
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.;
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000988; BAB66949.1; -.
```

DR GO: GO:0016021; C: integral to membrane; IEA.  
DR GO: GO:0005279; F: amino acid-polyamine transporter activity; IEA.  
DR GO: GO:0006865; P: amino acid transport; IEA.  
DR InterPro: IPR002293; AA/rel permease1.  
DR InterPro: IPR009045; Hedgehog\_sig\_N.  
KW Complete proteome; Hypothetical protein; Transmembrane.  
SQ SEQUENCE 530 AA; 59044 MW; A4B1AF2823DBE72C CRC64;

Query Match 60.0%; Score 39; DB 2; Length 530;  
Best Local Similarity 66.7%; Pred. No. 3.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 YYASSWPFAY 11  
| : |||||  
Db 117 YLAQWPFAY 125

## RESULT 40

Q6C9N1 Q6C9N1 PRELIMINARY; PRT; 534 AA.  
AC Q6C9N1;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Similar to tr|Q9URX1 Schizosaccharomyces pombe Hypothetical 55.8  
DE kDa.  
GN ORFNames=YALI0D098019;  
OS Yarrowia lipolytica CLIB99.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Dipodascaceae; Yarrowia.  
OX NCBI\_TaxID=284591;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CLIB99;  
RG Genolevures;  
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
RA Lafontaine I., de Montigny J., Marck C., Neveuglise C., Talla E.,  
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,  
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,  
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
RA Sweeney D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A., J.,  
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,  
RA Wincker P., Souciet J.L.;  
RT "Genome evolution in yeasts."  
RL Nature 430:35-44(2004).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CLIB99;  
RG Genoscope;  
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
RL ENBL; CR382130; CAG80819.1; -.  
DR InterPro: IPR010291; DUF895.  
DR Pfam: PF05978; DUF895; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 534 AA; 59157 MW; 3CEBDF1658D49222 CRC64;

Query Match 60.0%; Score 39; DB 2; Length 534;  
Best Local Similarity 55.6%; Pred. No. 3.4e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIYYASSWF 9  
| : |||||  
Db 292 PMFWASWNF 300

Search completed: June 3, 2005, 16:01:33  
Job time : 96.3333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 10:28:02 ; Search time 339.848 Seconds

(without alignments)  
2138.683 Million cell updates/sec

Title: US-10-089-452-24

Perfect score: 15

Sequence: 1 gacacctatgtgcac 15

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pi.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	15	6	AX113449 Sequence
2	15	100.0	15	6	AX113584 Sequence
3	15	100.0	15	6	AX816251 Sequence
4	15	100.0	360	6	AX113428 Sequence
5	15	100.0	360	6	AX113563 Sequence
6	15	100.0	360	6	AX816230 Sequence
7	15	100.0	53809	6	AX647061 Sequence
8	15	100.0	110000	2	EX649482_0
9	15	100.0	128361	9	AC004877 Homo sapi
10	15	100.0	144858	2	AC140972 Felis cat
11	15	100.0	145909	9	AC016482 Homo sapi
12	15	100.0	164834	2	AC136268 Rattus no
13	15	100.0	173370	2	AC116286 Rattus no
14	15	100.0	176742	9	AC007386 Homo sapi
15	15	100.0	192405	2	AC126160 Rattus no
16	15	100.0	196300	2	AC119907 Mus muscu
17	15	100.0	203623	10	AL929445 Mouse DNA
18	15	100.0	205507	2	AC134646 Rattus no
19	15	100.0	205588	9	AC068400 Homo sapi

20	15	100.0	211996	2	AC127651 Rattus no
c 21	15	100.0	215087	2	AC113594 Mus muscu
22	15	100.0	215875	2	AC104611 Rattus no
c 23	15	100.0	218858	2	AC103266 Rattus no
c 24	15	100.0	222535	10	AC125039 Mus muscu
25	15	100.0	225274	2	AC130251 Rattus no
26	15	100.0	229686	2	AC112734 Rattus no
27	15	100.0	230159	2	AC107472 Rattus no
c 28	15	100.0	230943	2	AC128807 Rattus no
29	15	100.0	231001	2	AC130981 Rattus no
30	15	100.0	232859	2	AC118408 Rattus no
31	15	100.0	236720	2	AC123083 Rattus no
c 32	15	100.0	242246	2	AC128408 Rattus no
c 33	15	100.0	253930	2	AC105621 Rattus no
34	15	100.0	266298	2	AC109096 Rattus no
c 35	14	93.3	185	6	AR449146 Sequence
c 36	14	93.3	201	11	BV203271 sqm21160
37	14	93.3	498	6	AX409913 Sequence
c 38	14	93.3	537	6	BD267157 Compositi
c 39	14	93.3	537	6	AR566678 Sequence
40	14	93.3	539	11	G73655 R2617R etio
41	14	93.3	599	11	BV159170 RPAMMSQ0
42	14	93.3	602	11	BV100061 RPAMMSQ0
43	14	93.3	602	11	BV163588 RPAMMSQ0
c 44	14	93.3	765	11	BV071186 S212P6120
c 45	14	93.3	848	6	BD267156 Compositi

#### ALIGNMENTS

RESULT 1  
LOCUS AX113449 15 bp DNA linear PAT 01-MAY-2001  
DEFINITION Sequence 24 from Patent WO0127612.  
ACCESSION AX113449  
VERSION AX113449.1 GI:13939708  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Reiter,C., Cullmann,G., Lakner,M., Truee,A., Dehnert,S. and Schwartz,G.  
TITLE Immuno-chromatographic rapid assay in order to detect acid-resistant microorganisms in the stool

JOURNAL Patent: WO 0127612-A 24 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung

mbH (DE) Location/Qualifiers

FEATURES  
source -  
1..15  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="CDR"

#### ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 1 GACACCTATGTGCAC 15

RESULT 2  
AX113584  
LOCUS AX113584 15 bp DNA linear PAT 01-MAY-2001  
DEFINITION Sequence 24 from Patent WO0127613.  
ACCESSION AX113584  
VERSION AX113584.1 GI:13939779  
KEYWORDS

SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Reiter,C., Cullmann,G., Heppner,P., Ringeis,A., Mueller,H. and Haindl,E.  
TITLE Improved method for the detection of acid resistant microorganisms in a stool  
JOURNAL Patent: WO 0127613-A 24 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung (DE)

FEATURES source  
Location/Qualifiers  
1..15  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="CDR"

ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 1 GACACCTATGTGCAC 15

RESULT 3  
AX816251  
LOCUS AX816251 15 bp DNA linear PAT 09-DEC-2003  
DEFINITION Sequence 24 from Patent EP1336850.  
ACCESSION AX816251  
VERSION AX816251.1 GI:39646773  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Reiter,C., Cullmann,G., Mueller,H., Heppner,P., Haindl,E. and Ringeis,A.  
TITLE Improved method for the detection of acid resistant microorganisms in a stool  
JOURNAL Patent: EP 1336850-A 24 20-AUG-2003;  
Connex Gesellschaft zur Optimierung von Forschung und Ent wicklung (DE)

FEATURES source  
Location/Qualifiers  
1..15  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 1 GACACCTATGTGCAC 15

RESULT 4  
AX113428  
LOCUS AX113428 360 bp DNA linear PAT 01-MAY-2001  
DEFINITION Sequence 3 from Patent WO0127612.  
ACCESSION AX113428  
VERSION AX113428.1 GI:13939696  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
AUTHORS Reiter,C., Cullmann,G., Lakner,M., Truse,A., Dehnert,S. and Schwartz,G.  
TITLE Immuno-chromatographic rapid assay in order to detect acid-resistant microorganisms in the stool  
JOURNAL Patent: WO 0127612-A 3 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung mbH (DE)

FEATURES source  
Location/Qualifiers  
1..360  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"

ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 360;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 91 GACACCTATGTGCAC 105

RESULT 5  
AX113563  
LOCUS AX113563 360 bp DNA linear PAT 01-MAY-2001  
DEFINITION Sequence 3 from Patent WO0127613.  
ACCESSION AX113563  
VERSION AX113563.1 GI:13939767  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
AUTHORS Reiter,C., Cullmann,G., Heppner,P., Ringeis,A., Mueller,H. and Haindl,E.  
TITLE Improved method for the detection of acid resistant microorganisms in a stool  
JOURNAL Patent: WO 0127613-A 3 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung (DE)

FEATURES source  
Location/Qualifiers  
1..360  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"

ORIGIN

Query Match 100.0%; Score 15; DB 6; Length 360;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 91 GACACCTATGTGCAC 105

RESULT 6  
AX816230  
LOCUS AX816230 360 bp DNA linear PAT 09-DEC-2003  
DEFINITION Sequence 3 from Patent EP1336850.  
ACCESSION AX816230  
VERSION AX816230.1 GI:39646761  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
AUTHORS Reiter,C., Cullmann,G., Mueller,H., Heppner,P., Haindl,E. and Ringeis,A.

**TITLE** Improved method for the detection of acid resistant microorganisms in a stool  
**JOURNAL** Patent: EP 1336850-A 3 20-AUG-2003;  
 Connex Gesellschaft zur Optimierung von Forschung und Ent wicklung (DE)

**FEATURES**  
 source  
 1. .360  
 /location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:10090"

**ORIGIN**

Query Match 100.0%; Score 15; DB 6; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
 |||||  
 Db 91 GACACCTATGTGCAC 105

**RESULT 7**  
**AX647061**  
**LOCUS** AX647061 53809 bp DNA linear PAT 04-MAR-2003  
**DEFINITION** Sequence 1253 from Patent EP1270724.  
**ACCESSION** AX647061  
**VERSION** AX647061.1 GI:28800002  
**KEYWORDS**  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1

**REFERENCE**  
**AUTHORS** Suwa,M., Asai,K., Akiyama,Y. and Aburatani,H.  
**TITLE** Guanosine triphosphate-binding protein coupled receptors  
**JOURNAL** Patent: EP 1270724-A 1253 02-JAN-2003;  
 National Institute of Advanced Industrial Science and Technology (JP) ; Center for Advanced Science and Technology Incubation, Ltd. (JP)

**FEATURES**  
 source

Location/Qualifiers  
 1. .53809  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 join(201. .288,1271. .1350,1650. .1677,3021. .3078,8331. .8353, 10809. .10841,12020. .12064,12426. .12862,13717. .13755, 19514. .19586,25189. .25368,27408. .27479,27653. .27826, 28136. .28300,30726. .30896,34508. .34678,36623. .36645, 37476. .37617,38132. .38331,39960. .40065,40143. .40361, 43119. .43365,43640. .43792,45218. .45540,45967. .46091, 48355. .48520,48693. .48820,49156. .49202,51179. .51256, 51673. .52253,53190. .53609)  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD69656.1"  
 /db\_xref="GI:28800003"

/translation="MGPEVLVIGQNVSVKQLVPEQGSLLHDDFMPEGGLAHLAA  
 TFGNSWRLPSEPDVGGVEACAGMAAGTVPRHCAPPVQSGSDGSGWGCACLEPASC  
 PCSLDTASSPLASAPGRGAAGVEGLTRALRVGGRSPAEDPARTLPAPPCGPFEP  
 RCSSGECTPRMWCQDQEDCADGSDRGCPAPHPAARGPHCVSPRLQCDGVRO  
 CPDGSDDPACVAPAPAMRPGGAGGTSTRAPSPPEAQCDCQPCDGRPG  
 QIPHSDQDPLLCPCAGAAALAGVAGAWMEAGWPCSVSCGGGHSQSCVDPDP  
 PKNGAPCPGASQBRAPGLQPCSGGTADRSQPQPPRAKRMGWRGVPAKVGWTFP  
 WMTSSCSQCLAPGGGGRSRLSCPSFGSDSGDGTQEEFSPVCPVFSIWGL  
 WAPWSTSCPDGGIQTRGRSRLAPGDDTTCQPHSQTRDCNTPCTVACGWSPTL  
 WSLSCSCNVGIRRRFRAGTAPPAAGGACQGTMEAEFCSLRPCPCGAGGWPWGPW  
 SHCSCSGGSLRSTRACDQPPQGLGDYCEGPRAQCEVQCALPCPCNCLAWREDLSP  
 GSVCPGSGVQCPTCGPLGSLQDGLCVPPHRCQYQPGAMAPVWQAFNCARCLT  
 ARTLGGALGALGTAVFRVAGSSCAPGAVLVLPAQLPARAAHRAHAEARHFVQ  
 AVATGEGDGGEGWAGEQDFQWREGLEARVHTGSLAPWTTASRPMVSVFGARFQ  
 RAPAPVGPAPGPMRWAAACOLGTVTPMPRTAGPRGATRMPATTAHAKLGSSPA  
 RLSPARLPVTPGATGPRGVPAATHAGPEGSRASGPPRRARGPQSVGRSSPRASPAL  
 SPRAPHCACALAPAPMGWTAGCSGASGAPAAAHLCPSVSHLCPVSVSHLPISGAPA

KAPHSCSPLLSCLLQASSWLSKAQPIILLPWRCPQVLSGCMGDTGLPOLAEGGHLVL  
 SPDPKGLSRATRHVLWALGCLLEQAGCCIPPEECNCWHTSAAGACWTLAPDGLQLG  
 CKWMSVSLGQGFQAVARPPVVGALWSDIGLVGVGWHHARPTQSGRSVTCSPA  
 LSAPLARCLVFPVPHASAGICSLVSPVSCRSPASLAVALGSGVHGGAFCSLVDVC  
 QGGDAGARELLSLPQGGSGVGNLVPQSPPPVPVSRVPTVSTSSVPTVPSPVSS  
 SCAPHSPLHPLHLCLPSSGLSPHLHLLCPPTVPSVPSCVLPQSPQSPSPVPSQSS  
 PPVSPQSPQSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP  
 HSPSLHLLCLCPQSPQSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP  
 PQSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP  
 PAPGASKKHORDLLPGFGRSELLQWVLP IQHPCHARGEFGTCGCGGNHVGQWPPPPA  
 ALSILGCGNQSP"

**ORIGIN**

Query Match 100.0%; Score 15; DB 6; Length 53809;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
 |||||  
 Db 4195 GACACCTATGTGCAC 4209

**RESULT 8**  
**AX649482\_0**  
**WPCOMMENT**

Sequence split into 4 fragments LOCUS BX649482 Accession BX649482

Fragment Name Begin End  
 BX649482\_0 1 110000  
 BX649482\_1 100001 210000  
 BX649482\_2 200001 310000  
 BX649482\_3 300001 358949

**LOCUS** BX649482 358949 bp DNA linear HTG 29-JAN-2004  
**DEFINITION** Danio rerio clone CH211-209H16, WORKING DRAFT SEQUENCE, 2 unordered pieces.

**ACCESSION** BX649482  
**VERSION** BX649482.6 GI:41392459  
**KEYWORDS** HTG; HTGS PHASE1; HTGS ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.  
**SOURCE** Danio rerio (zebrafish)  
**ORGANISM** Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 358949)  
 Phillimore,B.

**REFERENCE**  
**AUTHORS** Direct Submission  
**TITLE** Submitted (28-JAN-2004) Wellcome Trust Sanger Institute, Hinxton,  
**JOURNAL** Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Jan 29, 2004 this sequence version replaced gi:38141659.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC

Web site: http://www.sanger.ac.uk

Contact: zfish-help@sanger.ac.uk

----- Project Information

Center project name: zC209H16

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 166126 bases at least Q40

Consensus quality: 166133 bases at least Q30

Consensus quality: 166133 bases at least Q20

Insert size: 358849; sum-of-contigs

Insert size: 197037; 6.8% error; agarose-fp

Quality coverage: 3.46x in Q20 bases; sum-of-contigs Quality

coverage: 6.33x in Q20 bases; agarose-fp

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 2 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 166133: contig of 166133 bp in length  
 \* 166134 166233: gap of 100 bp  
 \* 166234 358949: contig of 192716 bp in length.

#### FEATURES

source  
 1. 358949  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="CH211-209H16"  
 /clone\_lib="CHORI-211"

#### misc\_feature

1. 166133  
 /note="assembly\_fragment:01905  
 clone end:SP6  
 vector side:left  
 clone\_end:T7  
 vector\_side:right"  
 166234..358949  
 /note="assembly\_fragment:01922"

#### misc\_feature

166234..358949

#### ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 110000;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACACCTATGTGCAC 15  
 |||||  
 Db 36253 GACACCTATGTGCAC 36267

RESULT 9  
 AC004877  
 LOCUS AC004877 128361 bp DNA linear PRI 27-JAN-2004  
 DEFINITION Homo sapiens PAC clone RP4-751H13 from 7, complete sequence.  
 AC004877  
 VERSION AC004877.1 GI:3638954  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1. (bases 1 to 128361)

Hallier,L.W., Fulton,R.S., Fulton,L.A., Graves,T.A., Pepin,K.H.,  
 Wagner-McPherson,C., Layman,D., Maas,J., Jaeger,S., Walker,R.,  
 Wylie,K., Sekhon,M., Becker,M.C., O'Laughlin,M.D., Schaller,M.E.,  
 Fewell,G.A., Delehaunty,K.D., Miner,T.L., Nash,W.E., Cordes,M.,  
 Du,H., Sun,H., Edwards,J., Bradshaw-Cordum,H., Ali,J., Andrews,S.,  
 Leack,A., Vanbrunt,A., Nguyen,C., Du,F., Lamar,B., Courtney,L.,  
 Kalicki,J., Ozersky,P., Bielicki,L., Scott,K., Holmes,A.,  
 Harkins,R., Harris,A., Strong,C.M., Hou,S., Tomlinson,C.,  
 Dauphin-Kohlberg,S., Kozlowicz-Reilly,A., Leonard,S., Rohlfing,T.,  
 Rock,S.M., Tin-Wollam,A.M., Abbott,A., Minx,P., Maupin,R.,  
 Strommatt,C., Latreille,P., Miller,N., Johnson,D., Murray,J.,  
 Woessner,J.P., Wendt,M.C., Yang,S.P., Schultz,B.R., Wallis,J.W.,  
 Spieth,J., Bieri,T.A., Nelson,J.O., Berkowicz,N., Wohlmann,P.E.,  
 Cook,L.L., Hickenbotham,M.T., Eldred,J., Williams,D., Bedell,J.A.,  
 Mardis,E.R., Clifton,S.W., Chisoe,S.L., Marra,M.A., Raymond,C.,  
 Haugen,B., Gillett,W., Zhou,Y., James,R., Phelps,K., Idanoto,S.,  
 Bubb,K., Simms,E., Levy,R., Clendinning,J., Kaul,R., Kent,W.J.,  
 Furey,T.S., Baertsch,R.A., Brent,M.R., Keibler,E., Flieck,P.,  
 Bork,P., Suyama,M., Bailey,J.A., Portnoy,M.E., Torrents,D.,  
 Chinwalla,A.T., Gish,W.R., Eddy,S.R., McPherson,J.D., Olson,M.V.,  
 Eichler,E.E., Green,E.D., Waterston,R.H. and Wilson,R.K.  
 The DNA sequence of human chromosome 7  
 Nature 424 (6945), 157-164 (2003)

#### TITLE

JOURNAL  
 MEDLINE  
 PUBMED  
 12853948

#### REFERENCE

2 (bases 1 to 128361)  
 Graves,T., Leonard,S. and Strommatt,C.  
 TITLE The sequence of Homo sapiens PAC clone RP4-751H13  
 JOURNAL Unpublished (2001)  
 3 (bases 1 to 128361)

#### AUTHORS

TITLE  
 JOURNAL

#### REFERENCE

AUTHORS  
 TITLE  
 JOURNAL

#### REFERENCE

AUTHORS  
 TITLE  
 JOURNAL

#### REFERENCE

AUTHORS  
 TITLE  
 JOURNAL

#### COMMENT

Waterston,R.H.  
 Direct Submission  
 Submitted (12-JUN-1998) Genome Sequencing Center, Washington  
 University, School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 4 (bases 1 to 128361)  
 Waterston,R.  
 Direct Submission  
 Submitted (19-SEP-1998) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 5 (bases 1 to 128361)  
 Waterston,R.  
 Direct Submission  
 Submitted (21-DEC-1999) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 6 (bases 1 to 128361)  
 Wilson,R.  
 Direct Submission  
 Submitted (27-JAN-2004) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Sep 19, 1998 this sequence version replaced gi:3213120.  
 ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu  
 Contact: sapiens@watson.wustl.edu  
 ----- Summary Statistics  
 -----  
 Center project name: H\_DJ0751H13  
 -----

NOTICE: This sequence may not represent the entire insert of this  
 clone. It may be shorter because we only sequence overlapping  
 clone sections once, or longer because we provide a small overlap  
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate  
 chemistry, or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by sequence  
 from more than one subclone; and the assembly was confirmed by  
 restriction digest.

#### MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and  
 sequencing collaboration between the NHGRI Chromosome 7 Mapping  
 Project (Eric D. Green, Director), John D. McPherson in the  
 Department of Genetics (Washington University), and the Washington  
 University Genome Sequencing Center. For additional information  
 about the map position of this sequence, see  
 http://www.nhgri.nih.gov/DIR/GRB/CHR7, send  
 mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu

#### SOURCE INFORMATION:

This clone was derived from human PAC library RPCI-4, prepared by  
 Pieter de Jong and coworkers at http://www.chori.org using the  
 method described by Ioannou et al., Nature Genetics 6:84-9 (1994).  
 The library is from one male donor.  
 The clone may be obtained either from Genome Systems, Inc.  
 (http://www.genomesystems.com) or Research Genetics, Inc.  
 (http://www.resgen.com); or from Pieter de Jong.  
 VECTOR: pCYPAC2

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP4-811N16, 200 bp overlap.  
 Actual start of this clone is at base position 1 of RP4-751H13  
 actual end is at 128361 of DJ0751H13.

#### Location/Qualifiers

1. 128361  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="7"

#### FEATURES

source

```
/map="7"
/clone="RP4-751HL3"
/clone_lib="RPC1-4"
20..159
/rpt_family="Alu"
860..941
/rpt_family="MER1_type"
2531..2885
/rpt_family="ERV1"
2924..3246
/rpt_family="Alu"
3300..3675
/rpt_family="ERV1"
4146..4287
/rpt_family="L2"
5253..5728
/rpt_family="MaLR"
5933..6243
/rpt_family="Alu"
7789..7827
/rpt_family="MIR"
8259..8311
/rpt_family="L2"
9773..10086
/rpt_family="Alu"
10297..10734
/rpt_family="MaLR"
11604..12221
/rpt_family="L1"
12222..12615
/rpt_family="ERV1"
12616..12752
/rpt_family="L1"
12768..12890
/rpt_family="MIR"
12891..13196
/rpt_family="Alu"
13197..13209
/rpt_family="MIR"
13314..13618
/rpt_family="Alu"
13636..13685
/rpt_family="L2"
13688..14004
/rpt_family="Alu"
14030..14162
/rpt_family="MIR"
14163..14192
/rpt_family="(TG)n"
15825..15931
/rpt_family="Alu"
16550..16702
/rpt_family="L1"
17763..17976
/rpt_family="MER1_type"
18090..19036
/rpt_family="L2"
19860..21428
/notes="CpG island (GC=70.4, o/e=0.73, #CpGs=171)"
22351..22486
/rpt_family="MIR"
22848..23146
/rpt_family="Alu"
23679..23794
/rpt_family="MIR"
23773..23846
/rpt_family="L2"
24894..24982
/rpt_family="L2"
24983..25271
/rpt_family="Alu"
25272..25615
/rpt_family="L2"

repeat_region 25616..25694
/rpt_family="GA-rich"
misc_feature 26933..29260
/notes="CpG island (GC=73.8, o/e=0.87, #CpGs=195)"
repeat_region 26960..27143
/rpt_family="C-rich"
repeat_region 27262..27404
/rpt_family="C-rich"
repeat_region 27427..27567
/rpt_family="C-rich"
repeat_region 27967..28307
/rpt_family="C-rich"
repeat_region 28409..28468
/rpt_family="GC-rich"

Query Match 100.0%; Score 15; DB 9; Length 128361;
Best Local Similarity 100.0%; Pred. NO. 2.4e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCAC 15
Db 37247 GACACCTATGTGCAC 37261

RESULT 10
AC140972 144858 bp DNA linear HTG 29-MAY-2003
LOCUS Felis catus clone RP86-144L3, WORKING DRAFT SEQUENCE, 11 ordered
DEFINITION pieces.
ACCESSION AC140972
VERSION AC140972.2 GI:31126657
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Felis catus
ORGANISM
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Carnivora; Fissipedia; Felidae; Felis.
REFERENCE 1 (bases 1 to 144858)
AUTHORS Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Carliaga,K.,
Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-O.,
Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
Maskeri,B., McDowell,J., Paguirigan,C., Pearson,R., Porcnoy,M.E.,
Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
Sison,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
TITLE NISC Comparative Sequencing Initiative
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 144858)
AUTHORS Green,E.D.
JOURNAL Direct Submission
TITLE Submitted (06-MAR-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 144858)
AUTHORS Green,E.D.
JOURNAL Direct Submission
TITLE Submitted (29-MAY-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT On May 29, 2003 this sequence version replaced gi:28867004.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: efa
Center clone name: 144L03

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
```

data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 142769 bases at least Q40  
Consensus quality: 143488 bases at least Q30  
Consensus quality: 143725 bases at least Q20  
Insert size: 168000; agarose-fp  
Insert size: 143858; sum-of-contigs  
Quality coverage: 10.45x in Q20 bases; agarose-fp  
Quality coverage: 12.20x in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and the accession number will be preserved.

- \* 1 908: contig of 908 bp in length
- \* 909 1008: gap of unknown length
- \* 1009 5899: contig of 4891 bp in length
- \* 5900 5999: gap of unknown length
- \* 6000 17014: contig of 11015 bp in length
- \* 17015 17114: gap of unknown length
- \* 17115 19195: contig of 2081 bp in length
- \* 19196 19295: gap of unknown length
- \* 19296 78377: contig of 58982 bp in length
- \* 78378 78377: gap of unknown length
- \* 85051 85051: contig of 6674 bp in length
- \* 85052 85151: gap of unknown length
- \* 85152 87453: contig of 2202 bp in length
- \* 87454 87453: gap of unknown length
- \* 87454 89562: contig of 2109 bp in length
- \* 89563 89662: gap of unknown length
- \* 89663 125084: contig of 35421 bp in length
- \* 125084 125183: gap of unknown length
- \* 125184 137615: contig of 12432 bp in length
- \* 137616 137716 144858: contig of 7143 bp in length.

FEATURES

source  
1. .144858  
/organism="Felis catus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9685"  
/clone="RP86-144L3"  
/clone\_lib="RP86"  
misc\_feature  
1. .908  
/notes="assembly\_fragment  
clone\_end:17  
vector\_side:left"  
misc\_feature  
1009. .5899  
/notes="assembly\_fragment"  
misc\_feature  
6000. .17014  
/notes="assembly\_fragment"  
misc\_feature  
17115. .19195  
/notes="assembly\_fragment"  
misc\_feature  
19296. .78277  
/notes="assembly\_fragment"  
misc\_feature  
78378. .85051  
/notes="assembly\_fragment"  
misc\_feature  
85152. .87353  
/notes="assembly\_fragment"

misc\_feature 87454. .89562  
/notes="assembly\_fragment"  
misc\_feature 89663. .125083  
/notes="assembly\_fragment"  
misc\_feature 125184. .137615  
/notes="assembly\_fragment"  
misc\_feature 137716. .144858  
/notes="assembly\_fragment"  
clone\_end:SP6  
vector\_side:right"

ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 144858;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACACCTATGTGCAC 15  
Db 111476 GACACCTATGTGCAC 111490

RESULT 11

AC016482  
LOCUS Homo sapiens chromosome 17, clone RP11-19G24, complete sequence.  
DEFINITION AC016482  
AC016482  
VERSION AC016482.5 GI:18464234  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 145909)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome 17, clone RP11-19G24  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 145909)  
REFERENCE 1 (bases 1 to 145909)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B., Brown, A., Castelle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (30-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

TITLE

JOURNAL  
REFERENCE  
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, A., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,



Strauss, N., Subramanian, A., Talamas, J., Tsefaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (08-FEB-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Feb 1, 2002 this sequence version replaced gi:16974276.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WtBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L916  
 Center clone name: 19\_G\_24  
 -----  
 Location/Qualifiers  
 1. 145909  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17"  
 /clone="RP11-19G24"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 1611..1800  
 /rpt\_family="CA-rich"  
 1864..2898  
 /rpt\_family="L1MC4"  
 2899..3189  
 /rpt\_family="AluSx"  
 3190..3386  
 /rpt\_family="L1MC4"  
 complement(3463..3668)  
 /rpt\_family="MIR"  
 /rpt\_family="MIR"  
 complement(4220..4285)  
 /rpt\_family="AluSx"  
 complement(4588..4661)  
 /rpt\_family="MIR"  
 5155..5185  
 /rpt\_family="AT-rich"  
 5727..5796  
 /rpt\_family="FLAM"  
 5797..5835  
 /rpt\_family="CA)n"  
 5840..6104  
 /rpt\_family="AluJo"  
 6106..6301  
 /rpt\_family="WER30"  
 complement(6424..6627)  
 /rpt\_family="MIR"  
 6744..7050  
 /rpt\_family="AluJo"  
 7423..7576  
 /rpt\_family="AluSg/x"  
 7688..7976  
 /rpt\_family="AluSp"  
 complement(10475..10769)  
 /rpt\_family="AluSx"  
 11017..11135  
 /rpt\_family="MIR3"  
 complement(11436..11536)  
 /rpt\_family="MIR"  
 12154..12207  
 /rpt\_family="A-rich"  
 14134..14167  
 /rpt\_family="(TCTA)n"

repeat\_region complement(14392..14606)  
 /rpt\_family="MIR"  
 15837..15921  
 /rpt\_family="L2"  
 16070..16292  
 /rpt\_family="AluJo"  
 16340..16633  
 /rpt\_family="AluJo"  
 complement(17872..17980)  
 /rpt\_family="MIR"  
 complement(18058..18366)  
 /rpt\_family="AluJb"  
 19446..19473  
 /rpt\_family="AT-rich"  
 20326..20520  
 /rpt\_family="L1MB4"  
 20715..20947  
 /rpt\_family="MIR"  
 21367..21472  
 /rpt\_family="L2"  
 21669..21709  
 /rpt\_family="(CA)n"  
 complement(21949..22259)  
 /rpt\_family="AluSx"  
 complement(22335..22631)  
 /rpt\_family="AluSx"  
 24068..24205  
 /rpt\_family="FRAM"  
 29646..29938  
 /rpt\_family="AluSx"  
 30016..30120  
 /rpt\_family="AluSp/q"  
 30311..30357  
 /rpt\_family="MIR"  
 32515..32658  
 /rpt\_family="L1PA4"  
 complement(32981..33054)  
 /rpt\_family="MADE1"  
 34588..34612  
 /rpt\_family="AT-rich"  
 37005..37119  
 /rpt\_family="GA-rich"  
 complement(37985..38116)  
 /rpt\_family="FLAM\_C"  
 38117..38139  
 /rpt\_family="AT-rich"  
 complement(38146..38730)  
 /rpt\_family="L1MB1"  
 38731..39040  
 /rpt\_family="AluSx"  
 complement(39041..39271)  
 /rpt\_family="L1MB1"  
 39511..39611  
 /rpt\_family="L2"  
 42074..42289  
 /rpt\_family="L3"  
 42351..42372  
 /rpt\_family="AT-rich"  
 42414..42456  
 /rpt\_family="AT-rich"  
 complement(43475..43627)  
 /rpt\_family="AluJo"  
 complement(43659..43741)  
 /rpt\_family="MIR"  
 complement(43759..44060)  
 /rpt\_family="AluSx"  
 44064..44094  
 /rpt\_family="(T)n"  
 complement(45076..45273)  
 /rpt\_family="MIR"  
 45277..45314  
 /rpt\_family="A-rich"  
 45765..45831

```

/rpt_family="CT-rich"
repeat_region 48762. .48814
/rpt_family="GA-rich"
repeat_region complement(50809. .50920)

```

Query Match 100.0%; Score 15; DB 9; Length 145909;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTCAC 15

Db 124135 GACACCTATGTCAC 124149

# RESULT 12

AC136268/c

LOCUS

DEFINITION Rattus norvegicus clone CH230-483P1, \*\*\* SEQUENCING IN PROGRESS

\*\*\*, 5 unordered pieces.

ACCESSION

AC136268.2 GI:25188356

VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_ENRICHED.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

1 (bases 1 to 164834)

Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anyaitebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Devila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flegg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Gunnaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensheewa, L., Loulseghe, H., Lozado, R. J., Lu, X., Ma, J.,

Mareshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,

Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervill, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nwaokelimeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,

Pasternak, S., Paul, H., Perez, A., Perez, L., Prannkoeh, C.,

Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,

Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,

Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,

Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,

Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,

Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,

Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,

Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,

Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,

Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Worley, K.,

## TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

## COMMENT

Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
Niederhauser, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,  
Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 164834)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (31-OCT-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 164834)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (23-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

On Nov 23, 2002 this sequence version replaced gi:24431539.

The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KCOQ

Center clone name: CH230-483P1

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 152353 bases at least Q40

Consensus quality: 154535 bases at least Q30

Consensus quality: 156110 bases at least Q20

Estimated insert size: 174847; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

----- NOTE: This is a 'working draft' sequence. It currently  
consists of 5 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

----- 1 21657: contig of 21657 bp in length

21658 21757: gap of unknown length

21758 101477: contig of 79720 bp in length

101478 101577: gap of unknown length

101578 158921: contig of 57344 bp in length

158922 159021: gap of unknown length

159022 160547: contig of 1526 bp in length

160548 160647: gap of unknown length

160648 164834: contig of 4187 bp in length.

Location/Qualifiers

1. .164834

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-483P1"

## FEATURES

source

## ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 164834;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCAC 15  
 |||||  
 Db 29670 GACACCTATGTGCAC 29656

RESULT 13  
 AC116286/c  
 LOCUS  
 DEFINITION Rattus norvegicus clone CH230-372F11, WORKING DRAFT SEQUENCE, 2  
 unordered pieces.  
 AC116286  
 AC116286.7 GI:25013250  
 VERSION HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
 KEYWORDS Rattus norvegicus (Norway rat)  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 173370)  
 Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
 Anyatebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
 Blawie, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flegg, N., Forbes, L., Foster, M., Foster, P.,  
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
 Gregeorge, E., Geer, K., Gill, R., Grady, M., Guetira, W., Guevara, W.,  
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,  
 Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,  
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorensuewa, L., Loulseghe, H., Lozado, R. J., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
 Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Nwaokemele, O., Okwunonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,  
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,  
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,  
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,  
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,  
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajda, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
 Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,  
 Williams, G., Willson, R., Wleczyk, R., Woodden, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,  
 Weinstock, G. and Gibbs, R. A.

TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 173370)  
 AUTHORS Worley, K. C.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-MAR-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 REFERENCE 3 (bases 1 to 173370)  
 AUTHORS Rat Genome Sequencing Consortium.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 COMMENT On Nov 15, 2002 this sequence version replaced gi:23617813.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GUDS  
 Center clone name: CH230-372F11  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 164395 bases at least Q40  
 Consensus quality: 165606 bases at least Q30  
 Consensus quality: 166400 bases at least Q20  
 Estimated insert size: 167474; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 172229: contig of 172229 bp in length  
 \* 172230 172329: gap of unknown length  
 \* 172330 173370: contig of 1041 bp in length.

FEATURES  
 source  
 1. 173370  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-372F11"  
 1. 2547  
 /notes="wgs\_end\_extension  
 clone\_end:T7"  
 3694-4596  
 /note="clone\_boundary  
 clone\_end:T7  
 site:  
 end sequence:BZ220031"  
 complement(164094..165012)  
 /notes="clone\_boundary  
 clone\_end:Sp6"

misc\_feature  
 misc\_feature  
 misc\_feature



```
repeat_region /note="match to EST N33519 (NID:g1153918) y43g12.s1"
3318.3461 /rpt_family="MER1_type"
repeat_region 3585.3708 /rpt_family="Alu"
misc_feature 3879.4088 /note="match to EST A1289177 (NID:g3932441) qn25f12.x1"
misc_feature 3879.4087 /note="match to EST H05038 (NID:g868590) y169a07.s1"
misc_feature 3881.4087 /note="match to EST AA939095 (NID:g3099008) oq22a03.s1"
misc_feature 3881.4072 /note="match to EST AA973070 (NID:g3148250) op31g08.s1"
misc_feature 3889.4087 /note="match to EST M86055 (NID:g274706)"
STS 3889.4002 /db_xref="GI:1340611"
misc_feature 4431.4590 /note="match to EST N43799 (NID:g1182327) y42g12.r1"
repeat_region 4683.4823 /rpt_family="MIR"
repeat_region 4964.5026 /rpt_family="MER103"
repeat_region 5088.5319 /rpt_family="MIR"
misc_feature 5389.5632 /note="match to EST AA939095 (NID:g3099008) oq22a03.s1"
misc_feature 5389.5455 /note="match to EST M86055 (NID:g274706)"
misc_feature 5389.5436 /note="match to EST H05038 (NID:g868590) y169a07.s1"
misc_feature 5633.6014 /note="match to EST AA701295 (NID:g2704460) z165c12.s1"
misc_feature 5653.5819 /note="match to EST H05038 (NID:g868590) y169a07.s1"
misc_feature 6097.6374 /note="match to EST H08919 (NID:g873741) y193a06.r1"
misc_feature 6100.6827 /note="match to EST AA203438 (NID:g1799149) zx56e06.r1"
misc_feature 6370.6824 /note="match to EST A1073503 (NID:g3400147) ool2e02.x1"
misc_feature 6407.6803 /note="match to EST R20076 (NID:g774710) yg39e03.r1"
misc_feature 6734.6915 /note="match to EST AA973070 (NID:g3148250) op31g08.s1"
misc_feature 7180.7668 /note="match to EST A1223034 (NID:g3805237) qg52h05.x1"
repeat_region 7438.7533 /rpt_family="MIR"
misc_feature 8588.8753 /note="match to EST AA973070 (NID:g3148250) op31g08.s1"
misc_feature 8679.8970 /note="match to EST AA443146 (NID:g2155821) zx69f10.r1"
repeat_region 9516.9740 /rpt_family="Alu"
misc_feature 9530.9611 /note="match to EST A1492056 (NID:g4393059) tg12a02.x1"
misc_feature 9540.9989 /note="match to EST R08010 (NID:g759933) yf17a05.s1"
misc_feature 10218.10603 /note="match to EST R08059 (NID:g759982) yf17a05.r1"
repeat_region 10880.10751 /rpt_family="MIR"
repeat_region 10986.11267 /rpt_family="Alu"
repeat_region 11387.11423 /rpt_family="MER2_type"
repeat_region 11424.11584 /rpt_family="Alu"
repeat_region 11585.11600 /rpt_family="Alu"
repeat_region 11601.11623 /rpt_family="Alu"

Query Match 100.0%; Score 15; DB 9; Length 176742;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCAC 15
|||||
Db 75051 GACACCTATGTGCAC 75037

RESULT 15
AC126160/c
LOCUS AC126160.3 GI:25074406
DEFINITION Rattus norvegicus clone CH230-513K6, WORKING DRAFT SEQUENCE.
ACCESSION AC126160
VERSION HTGS_PHASE2; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 192405)
Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Aryalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Deigado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Baves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falle,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,T., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzao,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,P.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartabeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,R., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlsczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

Direct Submission
TITLE
JOURNAL
Unpublished
```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 192405)
Worley, K. C.
Direct Submission
Submitted (04-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 192405)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:22855903.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAGX
Center clone name: CH230-513K6
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 172657 bases at least Q40
Consensus quality: 174253 bases at least Q30
Consensus quality: 175360 bases at least Q20
Estimated insert size: 178085; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as runs of N, however the sizes
of the gaps between them are based on estimates that have
provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
the accession number will be preserved.
*
* 1 192405: contig of 192405 bp in length.
FEATURES
source
Location/Qualifiers
1. .192405
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-513K6"
misc_feature
1. .2212
/notes="wgs contig"
misc_feature
6017. .7570
/notes="wgs contig"
misc_feature
complement(179873..180733)
/notes="clone boundary
clone_end:Sp6
site:
end sequence:RXBVJG3TV"
181620..183234
/notes="wgs end extension
clone_end:Sp6"
191335..192405
/notes="wgs_end_extension

```

```

ORIGIN
clone_end:Sp6"
Query Match 100.0%; Score 15; DB 2; Length 192405;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACACCTATGTGCAC 15
|||||
Db 20651 GACACCTATGTGCAC 20637
|||||
RESULT 16
AC119907
LOCUS
DEFINITION
AC119907 196300 bp DNA linear HTG 11-MAR-2004
Mus musculus chromosome 6 clone RP24-239G21 map 6, *** SEQUENCING
IN PROGRESS ***, 11 unordered pieces.
AC119907
AC119907.5 GI:453556279
HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 196300)
Birren, B., Nussbaum, C. and Lander, E.
Mus musculus chromosome 6, clone RP24-239G21
Unpublished
2 (bases 1 to 196300)
Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
McClean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (02-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 196300)
Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
Boquelavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,
Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagoopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,

```



SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 205507)
AUTHORS	Muzny,D.,Marle,M., Metzker,M.,Lee,S., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Ayoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Blawalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Georegeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokeleneh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,N., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 205507)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (28-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 205507)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Nov 20, 2002 this sequence version replaced gi:23343630. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.	
-----	Genome Center
Center:	Baylor College of Medicine
Center code:	BCM
Web site:	http://www.hgsc.bcm.tmc.edu/
Contact:	hgsc-help@bcm.tmc.edu
-----	Project Information
Center project name:	KBGM
Center clone name:	CH230-158A16
-----	Summary Statistics
Assembly program:	Phrap; version 0.990329
Consensus quality:	166420 bases at least Q40
Consensus quality:	170129 bases at least Q30
Consensus quality:	172883 bases at least Q20
Estimated insert size:	168437; sum-of-contigs estimation
Quality coverage:	6x in Q20 bases; sum-of-contigs estimation
-----	
* NOTE:	Estimated insert size may differ from sequence length
* (see	http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* consists of 3 contigs.	The true order of the pieces
* is not known and their order in this sequence record is	arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	be preserved.
* 1	153177: contig of 153177 bp in length
* 153178	153277: gap of unknown length
* 153278	204379: contig of 51102 bp in length
* 204380	204479: gap of unknown length
* 204480	205507: contig of 1028 bp in length.
FEATURES	Location/Qualifiers
source	1..205507
	/organism="Rattus norvegicus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10116"
	/clone="CH230-158A16"
misc_feature	1..1095
	/note="wgs contig"
misc_feature	3304..4924
	/note="wgs contig"
misc_feature	42660..43895
	/note="wgs contig"
misc_feature	170379..172454
	/note="wgs contig"
ORIGIN	
Query Match	100.0%; Score 15; DB 2; Length 205507;
Best Local Similarity	100.0%; Pred. No. 2.3e+02;
Matches 15; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 GACACCTATGTGCAC 15
Db	57967 GACACCTATGTGCAC 57953
RESULT 19	
AC068400/c	
LOCUS	AC068400 205588 bp DNA linear PRI 25-MAR-2003
DEFINITION	Homo sapiens chromosome 17, clone RP11-378E13, complete sequence.
ACCESSION	AC068400
VERSION	AC068400.10 GI:21281551
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)



**ORGANISM**

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**REFERENCE**

1 (bases 1 to 205588)

**AUTHORS**

Birren,B., Nusbaum,C. and Lander,E.

**TITLE**

Homo sapiens chromosome 17, clone RP11-378B13

**JOURNAL**

Unpublished

**AUTHORS**

2 (bases 1 to 205588)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,  
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,  
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,N.,  
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

**Direct Submission**

**Submitted (02-MAY-2000) Whitehead Institute/MIT Center for Genome Research,** 320 Charles Street, Cambridge, MA 02141, USA

**3 (bases 1 to 205588)**

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,  
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
Chazarob,B., Choepel,Y., Collangelo,M., Collins,S., Collamore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,  
Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,  
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

**Direct Submission**

**Submitted (03-MAY-2002) Whitehead Institute/MIT Center for Genome Research,** 320 Charles Street, Cambridge, MA 02141, USA

**4 (bases 1 to 205588)**

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,  
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
Chazarob,B., Choepel,Y., Collangelo,M., Collins,S., Collamore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,  
Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,  
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

**Direct Submission**

**Submitted (03-MAY-2002) Whitehead Institute/MIT Center for Genome Research,** 320 Charles Street, Cambridge, MA 02141, USA

**5 (bases 1 to 205588)**

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,  
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
Chazarob,B., Choepel,Y., Collangelo,M., Collins,S., Collamore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,  
Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,  
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

**Direct Submission**

**Submitted (31-MAY-2002) Whitehead Institute/MIT Center for Genome Research,** 320 Charles Street, Cambridge, MA 02141, USA

**6 (bases 1 to 205588)**

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,  
Collamore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,  
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,  
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,  
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,  
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,  
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,  
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,  
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

**Direct Submission**

**Submitted (25-MAR-2003) Whitehead Institute/MIT Center for Genome Research,** 320 Charles Street, Cambridge, MA 02141, USA

**On May 31, 2002 this sequence version replaced gi:20429394.**

**All repeats were identified using RepeatMasker:**

Smit,A.F.A. & Green,P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L10284  
Center clone name: 378 E 13

## FEATURES

```

source
1. .205588
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone_lib="RPCI-11 Human Male BAC"
33. .434
/rpt_family="L1PA15"
repeat_region
523. .692
/rpt_family="FRAM"
repeat_region
695. .840
/rpt_family="L1MC4"
repeat_region
877. .937
/rpt_family="MIR3"
repeat_region
1007. .1295
/rpt_family="AluSx"
repeat_region
complement(1437. .1728)
/rpt_family="AluSg"
repeat_region
complement(1993. .2302)
/rpt_family="AluY"
repeat_region
2502. .2800
/rpt_family="AluSx"
repeat_region
3039. .3227
/rpt_family="MIR"
repeat_region
3756. .4012
/rpt_family="AluJo"
repeat_region
complement(4022. .4354)
/rpt_family="AluSx"
repeat_region
complement(4436. .4746)
/rpt_family="AluSx"
repeat_region
4843. .5143
/rpt_family="AluSx"
repeat_region
complement(5485. .5645)
/rpt_family="MIR"
repeat_region
5681. .5979
/rpt_family="AluSp"
repeat_region
6045. .6336
/rpt_family="AluSg"
repeat_region
6343. .6572
/rpt_family="AluSx"
unsure
7037. .7041
/notes="<30 qual SNGL region"
repeat_region
7101. .7130
/rpt_family="GC-rich"

```

Query Match 100.0%; Score 15; DB 9; Length 205588;

Best Local Similarity 100.0%; Pred. NO. 2.3e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15

Db 195328 GACACCTATGTGCAC 195314

RESULT 20

AC127651

LOCUS AC127651 211996 bp DNA linear HTG 15-NOV-2002

DEFINITION Rattus norvegicus clone CH230-249E23, \*\*\* SEQUENCING IN PROGRESS

AC127651

VERSION AC127651.4 GI:25006957

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_ENRICHED.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 211996)

AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwakoeleneh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soza,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,R., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczkyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,P., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

## TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

## COMMENT

Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 211996)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 15, 2002 this sequence version replaced gi:23603352.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information

Center project name: GOHP  
Center clone name: CH230-249E23

----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 198403 bases at least Q40  
Consensus quality: 201136 bases at least Q30  
Consensus quality: 202737 bases at least Q20  
Estimated insert size: 206308; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 149338: contig of 149338 bp in length  
\* 149339 149438: gap of unknown length  
\* 149439 211996: contig of 62558 bp in length.

#### FEATURES

source Location/Qualifiers

1. 211996  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-249E23"  
7921..8741  
/note="clone boundary  
clone\_end:T7  
site:

end sequence:RXABI36TVB"  
complement(147545..148262)  
/note="clone boundary  
clone\_end:Sp6  
site:

end sequence:RXABI36TV"  
149439..151904  
/note="wgs end extension  
clone\_end:Sp6"

#### ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 211996;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GACACCTATGTCAC 15  
|||||  
Db 134714 GACACCTATGTCAC 134728

#### RESULT 21

AC113594/c  
LOCUS AC113594 215087 bp DNA linear HTG 10-MAR-2003  
DEFINITION Mus musculus clone RP23-365N23, WORKING DRAFT SEQUENCE, 8 unordered pieces.  
AC113594  
AC113594 GI:28894633  
VERSION HTG; HTGS PHASE1; HTGS DRAFT.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

#### REFERENCE

1 (bases 1 to 215087)  
Birren,B., Nusbaum,C. and Lander,E.  
Mus musculus, clone RP23-365N23  
Unpublished

2 (bases 1 to 215087)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T., Mlangwa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

#### Direct Submission

Submitted (03-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 215087)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouckgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fato,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hagos,B., Haggopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlangwa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

#### Direct Submission

Submitted (10-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 10, 2003 this sequence version replaced gi:28412058.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L23820

Center clone name: 365 N.23

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 213594 bases at least Q40

Consensus quality: 214029 bases at least Q30

Consensus quality: 214237 bases at least Q20

Insert size: 194000; agarose-fp  
 Insert size: 214387; sum-of-contigs  
 Quality coverage: 13.4 in Q20 bases; agarose-fp  
 Quality coverage: 12.1 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1372: contig of 1372 bp in length  
 \* 1373 1472: gap of 100 bp  
 \* 1473 4368: contig of 2896 bp in length  
 \* 4369 4468: gap of 100 bp  
 \* 4469 11913: contig of 7445 bp in length  
 \* 11914 12013: gap of 100 bp  
 \* 12014 22427: contig of 10414 bp in length  
 \* 22428 22527: gap of 100 bp  
 \* 22528 33152: contig of 10625 bp in length  
 \* 33153 33253: gap of 100 bp  
 \* 33253 49487: contig of 16235 bp in length  
 \* 49488 49588: gap of 100 bp  
 \* 49589 85653: contig of 36066 bp in length  
 \* 85654 85754: gap of 100 bp  
 \* 85754 215087: contig of 129334 bp in length.

## FEATURES

source  
 1. 215087  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="RP23-365N23"  
 /clone\_lib="RPCI-23 Female Mouse BAC"  
 1. 1372  
 /note="assembly\_fragment  
 clone\_end:SP6  
 vector\_side:left"  
 1473. 4368  
 /note="assembly\_fragment"  
 4469. 11913  
 /note="assembly\_fragment"  
 12014. 22427  
 /note="assembly\_fragment"  
 22528. 33152  
 /note="assembly\_fragment"  
 33253. 49487  
 /note="assembly\_fragment"  
 49588. 85653  
 /note="assembly\_fragment"  
 85754. 215087  
 /note="assembly\_fragment"

## ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 215087;  
 Best Local Similarity 100.0%; Pred. NO. 2.3e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTCAC 15  
 |||||  
 Db 163059 GACACCTATGTCAC 163045

## RESULT 22

AC104611  
 LOCUS AC104611 215875 bp DNA linear HTG 10-MAY-2003  
 DEFINITION Rattus norvegicus clone CH230-216C16, WORKING DRAFT SEQUENCE, 2  
 unordered pieces.  
 ACCESSION AC104611  
 VERSION AC104611.7 GI:30521101  
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
 SOURCE Rattus norvegicus (Norway rat)

## ORGANISM

REFERENCE  
AUTHORS

Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

1 (bases 1 to 215875)

Muzny,D.,Marie., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J.,  
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,  
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,  
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
 Blawie,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,  
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,  
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,  
 Chacko,J., Chavez,D., Chen,G., Chen,Z., Chen,Y., Chen,Z., Chu,J.,  
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cres,A., D'Souza,L.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,  
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,  
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,  
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,I., Fan,G.,  
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,  
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,  
 Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W.,  
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,  
 Harvey,Y., Hawlak,P., Hawes,A., Henderson,N., Hernandez,J.,  
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,  
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,  
 Karpach,S., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,  
 Karpach,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,  
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,  
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,  
 Lorensuhea,L., Loulseghe,H., Lozano,R.J., Lu,X., Ma,J.,  
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,  
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,  
 Mahoney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,  
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,  
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,  
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,  
 Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pai,S., Parks,K.,  
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,  
 Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.,-L.,  
 Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,  
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,P.,  
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,  
 Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,  
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,  
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,  
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,  
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,  
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,  
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,  
 Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,  
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,  
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von  
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
 Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 215875)

Worley,K.C.

Direct Submission

Submitted (14-DEC-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 215875)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:23681560.

The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
 in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GREK

Center clone name: CH230-216C16

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 202492 bases at least Q40

Consensus quality: 205070 bases at least Q30

Consensus quality: 206654 bases at least Q20

Estimated insert size: 212885; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 214054: contig of 214054 bp in length  
\* 214055 214154: gap of unknown length  
\* 214155 215875: contig of 1721 bp in length.

## FEATURES

```

source
1. .215875
   /organism="Rattus norvegicus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10116"
   /clone="CH230-216C16"
misc_feature
1224..1709
   /note="clone_boundary"
   /clone_end=T7
   site:EcoRI
end_sequence:BZ097326"
complement(213071..213760)
   /note="clone_boundary"
   /clone_end=Sp6
   site:EcoRI
end_sequence:BZ097327"

```

## ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 215875;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTCAC 15  
|||||  
Db 33055 GACACCTATGTCAC 33069

## RESULT 23

AC103266/c

LOCUS

DEFINITION AC103266 218958 bp DNA linear HTG 13-MAY-2003  
Rattus norvegicus clone CH230-173A15, \*\*\* SEQUENCING IN PROGRESS

ACCESSION AC103266

VERSION AC103266.5 GI:30578950

KEYWORDS HTG: HTGS PHASE1: HTGS\_DRAFT; HTGS\_ENRICHED.

SOURCE Rattus norvegicus (Norway rat)

## ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

1 (bases 1 to 218958)

## REFERENCE

## AUTHORS

Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alebrooks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,  
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorensuhewa, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,  
Millosevjevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L.,  
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,  
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,  
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajs, D.,  
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
Williams, G., Willson, R., Wleczyk, R., Woodden, H., Worley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 218958)

Worley, K.C.

Direct Submission

Submitted (24-NOV-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 218958)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:23269013.

The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: GJQS

Center clone name: CH230-173A15

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 189981 bases at least Q40

Consensus quality: 193948 bases at least Q30

Consensus quality: 196936 bases at least Q20

Estimated insert size: 197979; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 180406: contig of 180406 bp in length  
\* 180407 180506: gap of unknown length  
\* 180507 206026: contig of 25520 bp in length  
\* 206027 206126: gap of unknown length  
\* 206127 207180: contig of 1054 bp in length  
\* 207181 207280: gap of unknown length  
\* 207281 208411: contig of 1131 bp in length  
\* 208412 208511: gap of unknown length  
\* 208512 209617: contig of 1106 bp in length  
\* 209618 209717: gap of unknown length  
\* 209718 211921: contig of 2204 bp in length  
\* 211922 212021: gap of unknown length  
\* 212022 213341: contig of 1320 bp in length  
\* 213342 213441: gap of unknown length  
\* 213442 218858: contig of 5417 bp in length.

## FEATURES

## Source

1. 218858  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-173A15"  
1. 2537  
/note="wgs contig"  
192751..194185  
/note="wgs\_contig"

## ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 218858;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTCAC 15  
|||||

Db 147081 GACACCTATGTCAC 147067

RESULT 24

AC125039/c

LOCUS

222535 bp

DNA

linear

ROD 11-NOV-2003

## DEFINITION

Mus musculus BAC clone RP23-335C3 from chromosome 6, complete sequence.

ACCESSION

VERSION

AC125039.4 GI:25711049

KEYWORDS

SOURCE

ORGANISM

Mus musculus

(house mouse)

REFERENCE

1 (bases 1 to 222535)

Tomlinson, C., Bielicki, L., Schatzkamer, K. and Haakenson, W.

The sequence of Mus musculus BAC clone RP23-335C3

Unpublished (2001)

REFERENCE

2 (bases 1 to 222535)

Wilson, R.

Sequencing of Mus musculus

Unpublished (2001)

REFERENCE

3 (bases 1 to 222535)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (20-JUN-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

REFERENCE

4 (bases 1 to 222535)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (05-NOV-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

REFERENCE

5 (bases 1 to 222535)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (27-NOV-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

REFERENCE

6 (bases 1 to 222535)

Wilson, R.

Direct Submission

Submitted (11-NOV-2003) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Nov 27, 2002 this sequence version replaced gi:24580478.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)

----- Summary Statistics

Center project name: M\_BA0335C03

-----

NOTICE: This sequence may not represent the entire insert of this

clone. It may be shorter because we only sequence overlapping

clone sections once, or longer because we provide a small overlap

between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence

from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,

Department of Genetics, Washington University, St. Louis MO. For

additional information about the map position of this sequence, see

<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa

and Minako Tatenio in the laboratory of Pieter de Jong

(<http://www.chori.org>) from female C57BL/6J mouse kidney and/or

brain genomic DNA. The clone and detailed information can be

obtained from Research Genetics, Inc. (<http://www.resgen.com>) or

Pieter de Jong and coworkers at <http://www.chori.org>



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

# REFERENCE AUTHORS

1 (bases 1 to 225274)  
Muzny D, Marie, Metzker M, Lee, Abramson S, Adams C, Alder J, Allen C, Allen H, Albrooks S, Amin A, Anguiano D, Anyalebechi V, Aoyagi A, Ayodeji M, Baca E, Baden H, Baldwin D, Bandaranaike D, Barber M, Barnstead M, Benahmed F, Biewlo K, Blair J, Blankenburg K, Blyth P, Brown M, Bryant N, Buhay C, Burch P, Burrell K, Calderon E, Cardenas V, Carter K, Cavazos I, Ceasar H, Center A, Chacko J, Chavez D, Chen G, Chen R, Chen Z, Chu J, Cleveland C, Cockrell R, Cox C, Coyie M, Cree A, D'Souza L, Davila M, Davis C, Davy-Carroll L, De Ando C, Dederich D, Delgado O, Denison S, Deramo C, Ding Y, Dirh H, Divya K, Draper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K, Egan A, Escotto M, Eugene C, Evans C, Falls T, Fan G, Fernandez S, Finley M, Flagg N, Forbes L, Foster M, Foster P, Fraser C, Gabisi A, Ganta R, Garcia A, Garner T, Garza M, Gabregeorgis E, Geer K, Gill R, Grady M, Guerra W, Guevara W, Gunaratne P, Haaland W, Hamil C, Hamilton C, Hamilton K, Harvey Y, Havlak P, Hawes A, Henderson N, Hernandez J, Hernandez R, Hines S, Hladun S, Hodgson A, Hogues M, Hollins B, Howells S, Hulyk S, Hume J, Idlebird D, Jackson A, Jackson L, Jacob L, Jiang H, Johnson B, Johnson R, Jolivet A, Karpathy S, Kelly S, Kelly S, Khan Z, King L, Kovar C, Kowitz C, Kraft C, Lebow H, Levan J, Lewis L, Li Z, Liu J, Liu J, Liu Y, Liu Y, London P, Longacre S, Lopez J, Lorensuawa L, Loulsged H, Lozado R, Lu X, Ma J, Maheshwari M, Mahindartne M, Mahmoud M, Malloy K, Mangum A, Mangun B, Mapua P, Martin K, Martin R, Martinez B, Mathew S, McLeod M, McNeill T, Meenen E, Milosavljevic A, Miner G, Minja E, Montemayor J, Moore S, Morgan M, Morris K, Morris S, Munidasa M, Murphy M, Nair L, Nankervis C, Neal D, Newton N, Nguyen N, Norris S, Nwankwelen O, Okwunonu G, Olarnpunsagoon A, Pal S, Parks K, Patelnak S, Paul H, Perez A, Perez L, Pfannkuch C, Plopper F, Poinxter A, Popovic D, Primus E, Pu L, R, Puazo M, Quiroz J, Rachlin E, Reeves K, Regier M, Reigh R, Reilly B, Reilly M, Ren Y, Reuter M, Richards S, Riggs P, Rives C, Rodkey T, Rojas A, Rose M, Rose R, Ruiz S, Sanders W, Savery G, Scherer S, Scott G, Shatsman S, Shen H, Shetty J, Shvartsbeyn A, Sisson I, Sitter C, Smajd D, Sneed A, Sodergren E, Song X, Sorelle R, Sosa J, Steimle M, Strong R, Sutton A, Svatek A, Tabor P, Taylor C, Taylor T, Thomas R, Thomas S, Tingey A, Trejos Z, Usmani K, Valas R, Vera V, Villaseana D, Waldron L, Walker B, Wang J, Wang Q, Wang S, Warren J, Warren R, Wei X, White P, Williams G, Willson R, Wleczky R, Wooden H, Worley K, Wright D, Wright R, Wu J, Yakub S, Yen J, Yoon L, Yoon V, Yu F, Zhang J, Zhou J, Zhou X, Zhao S, Dunn D, von Niederhausern A, Weiss R, Smith D, Holt R, A., Smith H, O., Weinstock G, and Gibbs R, A.

## TITLE

## JOURNAL

Unpublished

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

Submitted (09-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:22855888. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: KANV  
Center clone name: CH230-157N23  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 211146 bases at least Q40  
Consensus quality: 213849 bases at least Q30  
Consensus quality: 215948 bases at least Q20  
Estimated insert size: 217419; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 223729: contig of 223729 bp in length  
\* 223730 223829: gap of unknown length  
\* 223830 225274: contig of 1445 bp in length.

## FEATURES

## source

1. 225274  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-157N23"

## misc\_feature

1. 1227  
/notes="wgs\_contig"

## misc\_feature

1278. 3304  
/notes="wgs contig"

## misc\_feature

222663. 223729  
/note="wgs\_contig"

## ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 225274;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15

Db 208615 GACACCTATGTGCAC 208629

|||||  
|||||

## RESULT 26

## AC112734

## LOCUS

## DEFINITION

AC112734 Rattus norvegicus clone CH230-52J2, WORKING DRAFT SEQUENCE, 4  
unordered pieces.  
AC112734 GI:30522136  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

## REFERENCE

1 (bases 1 to 229686)



## AUTHORS

Muzny D. Marie., Metzker M. Lee., Abramzon S., Adams C., Alder J., Allen C., Allen H., Albrooks S., Amin A., Anguiano D., Anyalbechi V., Aoyagi A., Ayodeji M., Baca E., Baden H., Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F., Biewald K., Blair J., Blankenburg K., Blyth P., Brown M., Bryant N., Buhay C., Burch P., Burrell K., Calderon E., Cardenas V., Carter K., Cavazos I., Ceasar H., Center A., Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J., Cleveland C., Cockrell R., Cox C., Coyte M., Cree A., D'Souza L., Davila M.B., Davis C., Davy-Carroll L., De Anda C., Dederich D., Delgado O., Denison S., Deramo C., Ding Y., Dinh H., Divya K., Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K., Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G., Fernandez S., Finley M., Flagg N., Forbes L., Foster P., Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M., Gebregiorgis E., Geer K., Gill R., Grady M., Guerra W., Guevara W., Gunaratne P., Haaland W., Hamill C., Hamilton C., Hamilton K., Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J., Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M., Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A., Jackson L., Jacob L., Jiang H., Johnson B., Johnson K., Jolivet A., Karpathy S., Kelly S., Kelly S., Khan Z., King L., Kovar C., Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J., Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J., Lorensuhewa L., Loulseghe H., Lozado R.J., Lu X., Ma J., Maheshwari M., Mahindaratne M., Mahmoud M., Malloy K., Mangum A., Mangum B., Mapua P., Martin K., Martin R., Martinez E., Mawhinney S., McLeod M.P., McNeill T.Z., Meenen E., Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S., Morgan M., Morris K., Morris S., Munidasa M., Murphy M., Nair L., Nankervis C., Neal D., Newton N., Nguyen N., Norris S., Nwaokaleme O., Okwuonu G., Olarnpunsagoon A., Pal S., Parks K., Patermak S., Paul H., Perez A., Perez L., Pfannkoch C., Plummer F., Poindexter A., Popovic D., Primus E., Pu L.-L., Puazo M., Quiroz J., Rachlin E., Reeves K., Regier M.A., Reigh R., Reilly B., Reilly M., Ren Y., Reuter M., Richards S., Riggs F., Rives C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.J., Sanders W., Savery G., Scherer S., Scott G., Shatsman S., Shen H., Shetty J., Shvartsbeyn A., Sisson I., Sitter C.D., Smajes D., Sneed A., Sodergren E., Song X.-Z., Sorelle R., Sosa J., Steimle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C., Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Usmani K., Valas R., Vera V., Villalana D., Waldron L., Walker B., Wang J., Wang Q., Wang S., Warren J., Warren R., Wei X., White F., Williams G., Willson R., Wleczyk R., Woodson H., Worley K., Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V., Yu F., Zhang J., Zhou J., Zhou X., Zhao S., Dunn D., von Niederhausern A., Weiss R., Smith D.R., Holt R.A., Smith H.O., Weinstock G. and Gibbs R.A.

## TITLE

## JOURNAL

REFERENCE  
2 (bases 1 to 229686)

## AUTHORS

## TITLE

Submitted (24-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

## JOURNAL

3 (bases 1 to 229686)

## REFERENCE

## TITLE

## JOURNAL

## COMMENT

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 10, 2003 this sequence version replaced gi:23664966.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu

----- Project Information  
Center project name: GGYE  
Center clone name: CH230-52J2  
----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 213435 bases at least Q40

Consensus quality: 215611 bases at least Q30

Consensus quality: 216973 bases at least Q20

Estimated insert size: 223908; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 4 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 226140: contig of 226140 bp in length

\* 226141 226240: gap of unknown length

\* 226241 227278: contig of 1038 bp in length

\* 227279 227378: gap of unknown length

\* 227379 228509: contig of 1131 bp in length

\* 228510 228609: gap of unknown length

\* 228610 229686: contig of 1077 bp in length.

\* Location/Qualifiers

1. 229686

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-52J2"

1. 1373

/note="wgs contig"

5707. 7072

/note="wgs contig"

complement(225343. 226141)

/note="clone boundary"

clone end:T7

site:ECORI

end\_sequence:BH315923"

## ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 229686;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15

|||||

Db 17140 GACACCTATGTGCAC 17154

|||||

RESULT 27

AC107472

LOCUS

DEFINITION

AC107472

ACCSSTON

VERSION

AC107472.5 GI:30580650

KEYWORDS

HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_ENRICHED.

SOURCE

Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 230159)

REFERENCE  
AUTHORS  
Muzny, D., Marle, Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, S., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levay, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenuhewa, L., Loulsegged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeleneh, O., Okunolu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinkle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willison, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission  
Unpublished  
2 (bases 1 to 230159)  
Worley, K. C.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (22-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 230159)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
On May 13, 2003 this sequence version replaced gi:23269081. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GKEM  
Center clone name: CH230-194N20

----- Summary Statistics  
Assembly program: Atlas 3.0;  
Consensus quality: 202289 bases at least Q40  
Consensus quality: 207215 bases at least Q30  
Consensus quality: 209951 bases at least Q20  
Estimated insert size: 215253; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
-----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 93296: contig of 93296 bp in length  
\* 93297 93396: gap of unknown length  
\* 93397 167763: contig of 74367 bp in length  
\* 167764 167863: gap of unknown length  
\* 167864 182020: contig of 14157 bp in length  
\* 182021 182120: gap of unknown length  
\* 182121 227247: contig of 45127 bp in length  
\* 227248 227347: gap of unknown length  
\* 227348 228792: contig of 1445 bp in length  
\* 228793 228892: gap of unknown length  
\* 228893 230159: contig of 1267 bp in length.

## FEATURES

Location/Qualifiers  
source  
1. .230159  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-194N20"  
93397. .95006  
misc\_feature  
/note="wgs\_contig"

## ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 230159;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCAC 15

Db 228492 GACACCTATGTGCAC 228506

## RESULT 28

AC128807/c  
LOCUS 230943 bp DNA linear HTG 20-NOV-2002  
DEFINITION Rattus norvegicus clone CH230-14211, WORKING DRAFT SEQUENCE, 3  
unordered pieces.  
AC128807  
AC128807.3 GI:25139596  
HTG: HTGS\_PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 230943)  
Muzny,D.,Marle., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,  
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,  
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,  
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,  
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,  
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,  
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,X., Chen,Z., Chu,J.,  
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,  
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,  
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,  
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,  
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,  
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,  
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,  
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W.,  
Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,  
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,  
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,  
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,  
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,  
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,  
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,  
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,  
Lorensuhewa,L., Loulseghe,H., Lozada,R.J., Lu,X., Ma,J.,  
Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,  
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,  
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,  
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,  
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,  
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,  
Nwankweli,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,  
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,  
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,  
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,  
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,  
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,  
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,  
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smaje,D.,  
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,  
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,  
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,  
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,  
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,  
Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K.,  
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,  
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von  
Niederhausern,A., Weises,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
Weinstock,G. and Gibbs,R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 230943)  
Worley,K.C.  
Direct Submission  
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 230943)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Nov 20, 2002 this sequence version replaced gi:23269430.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: KMM  
Center clone name: CH230-14211  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 209176 bases at least Q40  
Consensus quality: 210687 bases at least Q30  
Consensus quality: 211742 bases at least Q20  
Estimated insert size: 216317; sum-of-contigs estimation  
Quality coverage: 11x in Q20 bases; sum-of-contigs estimation  
-----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 18108: contig of 18108 bp in length  
\* 18109 18208: gap of unknown length  
\* 109809 109880: contig of 91672 bp in length  
\* 109881 109980: gap of unknown length  
\* 109981 230943: contig of 120963 bp in length.

FEATURES  
source

1. 230943  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-14211"  
misc\_feature 18209..19465  
misc\_feature 109981..112198  
/notes="wgs\_contig"

## ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 230943;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCAC 15

Db 197406 GACACCTATGTGCAC 197392

## RESULT 29

AC130981  
LOCUS 231001 bp DNA linear HTG 13-MAY-2003  
DEFINITION Rattus norvegicus clone CH230-229C14, WORKING DRAFT SEQUENCE.  
ACCESSION AC130981  
VERSION AC130981.3 GI:30578466  
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS\_FULLTOP.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus

## REFERENCE

1 (bases 1 to 231001)  
Muzny,D.,Marle., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,  
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabris, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregregios, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okunolu, G., Olarunpasegun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajda, D., Steed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlecyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

## Direct Submission

Unpublished

2 (bases 1 to 231001)  
Rat Genome Sequencing Consortium.

Direct Submission

Submitted (16-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 231001)

Rat Genome Sequencing Consortium.

## Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:23664662.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature

## table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GJYQ
Center clone name: CH230-229C14
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 216581 bases at least Q40
Consensus quality: 218902 bases at least Q30
Consensus quality: 220353 bases at least Q20
Estimated insert size: 225199; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 231001: contig of 231001 bp in length.
----- Location/Qualifiers
source
    1..231001
       /organism="Rattus norvegicus"
       /mol_type="genomic DNA"
       /db_xref="taxon:10116"
       /clone="CH230-229C14"
misc_feature
    1..1186
       /notes="wgs_contig"
       /complement(229055..229809)
       /notes="clone boundary"
       clone_end:Sp6
       site:ECORI
       end_sequence:B2108940"
ORIGIN
Query Match      100.0%; Score 15; DB 2; Length 231001;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACACTATGTGCAC 15
Db      83317 GACACTATGTGCAC 83331
|||||
DEFINITION
Rattus norvegicus clone CH230-213D23, *** SEQUENCING IN PROGRESS
AC118408
AC118408.6 GI:25138166
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 232859)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

```

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, W., Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowitz, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, L., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mepua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaekeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindester, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J. S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shivartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

## TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

## COMMENT

Submitted (17-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 232859)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On Nov 20, 2002 this sequence version replaced gi:23812546.  
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GUGW

Center clone name: CH230-213D23

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 222382 bases at least Q40

Consensus quality: 224780 bases at least Q30

Consensus quality: 226432 bases at least Q20

Estimated insert size: 228853; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 1 contigs. Gaps between the contigs

\* are represented as runs of N. The order of the pieces

\* is believed to be correct as given, however the sizes

\* of the gaps between them are based on estimates that have

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 232859: contig of 232859 bp in length.

-----

Location/Qualifiers

1. 232859

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clones="CH230-213D23"

1. 3704

/note="wgs end extension

clone\_end:Sp6"

3755. 5001

/note="wgs end extension

clone\_end:Sp6"

7827. 8701

/note="clone boundary

clone\_end:Sp6

site:

end sequence:RWBLK24TVB"

complement(229193..230066)

/note="clone boundary

clone\_end:T7

site:

end sequence:RWBLK24TJC"

230854..232859

/note="wgs end extension

clone\_end:T7"

-----

misc\_feature

-----

misc\_feature

-----

misc\_feature

-----

misc\_feature

-----

misc\_feature

-----

## ORIGIN

Query Match 100.0%; Score 15; DB 2; Length 232859;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15

Db 56362 GACACCTATGTGCAC 56376

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

## RESULT 31

AC123083

LOCUS

DEFINITION

AC123083

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AC123083

LOCUS

DEFINITION

AC123083

AC123083 Rattus norvegicus clone CH230-10J13, \*\*\* SEQUENCING IN PROGRESS

\*\*\*, 3 unordered pieces.

AC123083 AC123083.4 GI:24818090

HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_ENRICHED.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE  
AUTHORS

1 (bases 1 to 236720)  
Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Ding, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Gunaratne, P., Healand, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, E., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowitz, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, L., Loulseg, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, B., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankweli, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Riley, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajd, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 236720)  
Worley, K.C.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL

Direct Submission  
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 236720)  
Rat Genome Sequencing Consortium.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL

Direct Submission  
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 9, 2002 this sequence version replaced gi:23264486.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GDWP  
Center clone name: CH230-10J13  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 217461 bases at least Q40  
Consensus quality: 221406 bases at least Q30  
Consensus quality: 224047 bases at least Q20  
Estimated insert size: 222742; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.  
\* 1 234147: contig of 234147 bp in length  
\* 234148 234247: gap of unknown length  
\* 234248 235526: contig of 1279 bp in length  
\* 235527 235626: gap of unknown length  
\* 235627 236720: contig of 1094 bp in length.  
----- Location/Qualifiers  
source  
1. .236720  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-10J13"  
1. .1120  
/notes="wgs contig"  
2009. .3235  
/notes="wgs contig"  
/notes="wgs contig"  
214231. .215526  
/notes="wgs contig"  
complement(232514. .233171)  
/notes="clone boundary"  
clone\_end:T7  
site:EcoRI  
end\_sequence:BH306834"

misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature

Query Match 100.0%; Score 15; DB 2; Length 236720;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCAC 15  
|||||  
DB 125982 GACACCTATGTGCAC 125996

RESULT 32  
AC128408/c  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-156B8, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 4 unordered pieces.  
AC128408  
AC128408.3 GI:25138158  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
Rattus norvegicus (Norway rat)

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GYU  
Center clone name: CH230-156B8  
----- Summary Statistics  
Assembly program: Phrap; version 0.90329  
Consensus quality: 137461 bases at least Q40  
Consensus quality: 200807 bases at least Q30  
Consensus quality: 203054 bases at least Q20  
Estimated insert size: 201964; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 36980: contig of 36980 bp in length  
36981 37080: gap of unknown length  
37081 239359: contig of 202279 bp in length  
239360 239459: gap of unknown length  
239460 240897: contig of 1438 bp in length  
240898 240997: gap of unknown length  
240998 242246: contig of 1249 bp in length.

FEATURES  
source  
1. 242246  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-156B8"  
/notes="complement(36484. 36672)  
clone\_end:Spf  
site:  
end\_sequence:BH365639"  
135000..136226  
/notes="wgs contig"  
194888..194929  
/notes="clone boundary  
clone\_end:T7  
site:  
end\_sequence:BH365638"

misc\_feature  
misc\_feature  
misc\_feature

misc\_feature  
misc\_feature  
misc\_feature

misc\_feature  
misc\_feature  
misc\_feature

ORIGIN  
Query Match 100.0%; Score 15; DB 2; Length 242246;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 3701 GACACCTATGTGCAC 3687

RESULT 33  
AC105621/c  
LOCUS AC105621 253930 bp DNA linear HTG 19-NOV-2002



DEFINITION	Rattus norvegicus clone CH230-113H3, WORKING DRAFT SEQUENCE.	COMMENT	On Nov 19, 2002 this sequence version replaced gi:23664549. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be whole contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
ACCESSION	AC105621		
VERSION	AC105621.4 GI:25075212		
KEYWORDS	HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 253930)		
AUTHORS	Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregeorge, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, S., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Loulsegged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakeleneh, O., Okunnu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.		
	Direct Submission		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 253930)		
AUTHORS	Worley, K. C.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
	3 (bases 1 to 253930)		
REFERENCE	Rat Genome Sequencing Consortium.		
AUTHORS	Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
TITLE	Direct Submission		
JOURNAL	Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
		FEATURES	
		source	
			1 253930: contig of 253930 bp in length.
			Location/Qualifiers
			1. .253930
			/organism="Rattus norvegicus"
			/mol_type="genomic DNA"
			/db_xref="taxon:10116"
			/clone="CH230-113H3"
		misc_feature	1. .1784
			/note="wgs_end_extension"
			clone_end:T7"
			3955. .4825
			/note="clone boundary"
			clone_end:T7"
			site:
			end_sequence:BH261574"
			3990. .4718
			/note="clone boundary"
			clone_end:T7"
			site:
			end_sequence:BH261574"
			167592. .170065
			/note="wgs_end_extension"
			clone_end:T7"
			170116. .171920
			/note="wgs_end_extension"
			clone_end:T7"
			191435. .191485
			/note="clone boundary"
			clone_end:Sp6
			site:
			end_sequence:BH261576"
			252536. .253930
			/note="wgs_end_extension"



```

ORIGIN
Query Match 100.0%; Score 15; DB 2; Length 253930;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTCAC 15
Db 168192 GACACCTATGTCAC 168178

RESULT 34
AC109096
LOCUS AC109096 266298 bp DNA linear HTG 09-OCT-2002
DEFINITION Rattus norvegicus clone CH230-22206, *** SEQUENCING IN PROGRESS
ACCION0906
VERSION AC109096.4 GI:23195469
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 266298)
Muzny,D.,Marle., Metzker,M.,Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biewalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,X., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowig,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwakoleneh,O., Okwunou,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von

```

Niederhausem,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
 Weinstock,G. and Gibbs,R.A.  
 Unpublished  
 2 (bases 1 to 266298)  
 Worley,K.C.  
 Direct Submission  
 Submitted (03-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 266298)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (09-OCT-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Sep 19, 2002 this sequence version replaced gi:21738014.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GQCS  
 Center clone name: CH230-22206  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 227831 bases at least Q40  
 Consensus quality: 231515 bases at least Q30  
 Consensus quality: 233316 bases at least Q20  
 Estimated insert size: 250977; sum-of-contigs estimation  
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

-----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)  
 \* NOTE: This sequence may represent more than one clone  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 254250: contig of 254250 bp in length  
 \* 254251 254350: gap of unknown length  
 \* 254351 266298: contig of 11948 bp in length.

FEATURES  
 source  
 1. 266298  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-22206"  
 1. 1141  
 /notes="wgs contig"  
 105672. 106717  
 /notes="wgs contig"  
 134904. 135946  
 /notes="wgs contig"  
 complement(252617..253472)  
 /note="clone\_boundary"

```

clone_end:Sp6
site:ECORI
end_sequence:RWBMV87TV"
254351..257176
/misc_feature
/clone_lib="Human DNA (Sequenom)"
/clone_end:Sp6"

ORIGIN
Query Match 100.0%; Score 15; DB 2; Length 266298;
Best Local Similarity 100.0%; Pred. No. 2.3e+02; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;

Qy 1 GACACCTATGTGCAC 15
|||||
Db 37281 GACACCTATGTGCAC 37295

RESULT 35
AR449146/c
LOCUS AR449146 185 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 38 from patent US 6673909.
ACCESSION AR449146
VERSION AR449146.1 GI:42678197
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 185)
AUTHORS Brown,R.H. Jr., Liu,J., Roki,M., Ho,M.F. and Matsuda-Asada,C.
TITLE Oligonucleotides for dysferlin, a gene mutated in distal myopathy
JOURNAL Patent: US 6673909-A 38 06-JAN-2004;
FEATURES
source 1..185
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 93.3%; Score 14; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.5e+03; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0;

Qy 1 GACACCTATGTGCA 14
|||||
Db 176 GACACCTATGTGCA 163

RESULT 36
BV203271/c
LOCUS BV203271 201 bp DNA linear STS 10-JUN-2004
DEFINITION sqm211607 Human DNA (Sequenom) Homo sapiens STS genomic, sequence tagged site.
ACCESSION BV203271
VERSION BV203271.1 GI:48172604
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 201)
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M., Cantor,C.R. and Braun,A.
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Andreas Braun
Pharmaceuticals division
Sequenom, Inc.
3595 John Hopkins Court, San Diego, CA 92121, USA
Tel: 18582029018
Fax: 18582029020
Email: abraun@sequenom.com

```

```

Primer A: No primer sequence submitted
Primer B: No primer sequence submitted
STS size: 201.
Location/Qualifiers
1..201
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human DNA (Sequenom)"
<1..>201

ORIGIN
Query Match 93.3%; Score 14; DB 11; Length 201;
Best Local Similarity 100.0%; Pred. No. 1.5e+03; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0;

Qy 2 ACACCTATGTGCAC 15
|||||
Db 152 ACACCTATGTGCAC 139

RESULT 37
AX409913
LOCUS AX409913 498 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2560 from Patent WO0229103.
ACCESSION AX409913
VERSION AX409913.1 GI:21442618
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2560 11-APR-2002;
FEATURES
source 1..498
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. N53757"

ORIGIN
Query Match 93.3%; Score 14; DB 6; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.4e+03; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0;

Qy 2 ACACCTATGTGCAC 15
|||||
Db 36 ACACCTATGTGCAC 49

RESULT 38
BD267157/c
LOCUS BD267157 537 bp DNA linear PAT 17-JUL-2003
DEFINITION Compositions isolated from plant cells and utilization of the same in modifying plant cell signal transduction.
ACCESSION BD267157
VERSION BD267157.1 GI:33076925
KEYWORDS JP 2002534115-A/236.
SOURCE Eucalyptus grandis
ORGANISM Eucalyptus grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Myrtales; Myrtaceae; Eucalyptus.
REFERENCE 1 (bases 1 to 537)
AUTHORS Strabala,T.J. and Nieuwenhuizen,N.J.
TITLE Compositions isolated from plant cells and utilization of the same in modifying plant cell signal transduction
JOURNAL Patent: JP 2002534115-A 236 15-OCT-2002;
GENESIS RESEARCH AND DEVELOPMENT CORP LTD

```

COMMENT OS Eucalyptus grandis (flooded gum)  
 PN JP 2002534115-A/236  
 PD 15-OCT-2002  
 PF 11-JAN-2000 JP 2000593728  
 PR 12-JAN-1999 US 09/228986.01-NOV-1999 US 60/162866 PI  
 TIMOTHY J STRABALA,NICOLAAS J NIEUWENHUIZEN  
 PC C12N15/09,A01H5/00,C07K14/415,C07K19/00,C12N5/10,C12N9/12// PC  
 (C12N9/12,C12R1:91),C12N15/00,C12N5/00  
 CC Compositions isolated from plant cells and utilization of the  
 same in  
 CC modifying plant cell signal transduction  
 CC Key Location/Qualifiers  
 FT source 1. 537  
 FT /organism='Eucalyptus grandis (flooded gum)'.  
 FT Location/Qualifiers

FEATURES  
 source  
 1. .537  
 /organism="Eucalyptus grandis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:71139"

## ORIGIN

Query Match 93.3%; Score 14; DB 6; Length 537;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 2 ACACCTATGTGCAC 15  
 |||||  
 Db 156 ACACCTATGTGCAC 143

RESULT 19  
 AR566678/c  
 LOCUS AR566678 537 bp DNA linear PAT 08-OCT-2004  
 DEFINITION Sequence 299 from patent US 6768041.  
 ACCESSION AR566678  
 VERSION AR566678.1 GI:53983713  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 537)  
 AUTHORS Strabala,T. and Nieuwenhuizen,N.  
 TITLE Compositions isolated from plant cells and their use in the  
 modification of plant cell signaling  
 JOURNAL Patent: US 6768041-A 299 27-JUL-2004;  
 FEATURES Location/Qualifiers  
 source  
 1. .537  
 /organism="unknown"  
 /mol\_type="genomic DNA"

## ORIGIN

Query Match 93.3%; Score 14; DB 6; Length 537;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 2 ACACCTATGTGCAC 15  
 |||||  
 Db 156 ACACCTATGTGCAC 143

RESULT 40  
 G73655  
 LOCUS G73655 539 bp DNA linear STS 16-JUL-2002  
 DEFINITION RZ617R etiolated leaf tissue of rice Oryza sativa STS genomic clone  
 RZ617 sequencing direction=reverse, sequence tagged site.  
 ACCESSION G73655  
 VERSION G73655.1 GI:19697250  
 KEYWORDS STS.  
 SOURCE Oryza sativa (indica cultivar-group)  
 ORGANISM Oryza sativa (indica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

## REFERENCE

1 (bases 1 to 539)  
 McCouch,S.R.  
 Oryza sativa STS  
 Unpublished (2002)  
 CONTACT: Susan R. McCouch  
 Cornell University  
 Dept. of Plant Breeding, Ithaca, NY 14853-1901, USA  
 Tel: 6072550420  
 Fax: 6072556683  
 Email: srm4@cornell.edu  
 Primer A: M13 universal Forward GTAAACGACGGCCAGT  
 Primer B: M13 Universal Reverse AACAGCTATGACCATG  
 STS size: 539  
 Protocol:

Template: 20-100ng  
 Primer: 5pmol each  
 DNTps: 40nmol  
 Taq polymerase: 5units  
 Total volume: 50ul

Buffer:  
 Tris-HCl: 100mM  
 KCl: 500mM  
 MgCl2: 15mM  
 Gelatin: 0.1%  
 PH: 8.3

This is a partial sequence of the RFLP clone named above that was mapped at Cornell University Plant Breeding Dept. and sequenced at the Genome Sequencing Center at Cold Spring Harbor Laboratory. This marker is located on rice chromosome 8. For citations and other related information concerning this probe, please refer to the Gramene database at http.

## FEATURES

source  
 1. .539  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /cultivar="IR36"  
 /db\_xref="taxon:39946"  
 /map="8"  
 /clone="RZ617"  
 /note="Vector: Lambda ZAP II/pBluescript; V-type: Plasmid;  
 A Lambda ZAP II cDNA library was constructed from mRNA  
 extracted from etiolated leaf tissue of the rice cultivar  
 'IR36' and converted to pBluescript (amp resistant) as  
 described in Causse et al. (1994) Genetics 138:1251-1274.  
 For insert amplification, Use M13 forward and reverse  
 primers. Restriction site is SmaI. Clones from this  
 library are designated with the prefix 'RZ'."

## ORIGIN

Query Match 93.3%; Score 14; DB 11; Length 539;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GACACCTATGTGCA 14  
 |||||  
 Db 441 GACACCTATGTGCA 454

Search completed: June 4, 2005, 14:31:35  
 Job time : 352.848 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 10:23:32 ; Search time 92.4242 Seconds  
(without alignments)  
960.745 Million cell updates/sec

Title: US-10-089-452-24

Perfect score: 15

Sequence: 1 gacacctatgtgcac 15

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseqn16Dec04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	15	4	Aaf88108 H. pylori
2	15	100.0	15	4	Aaf88051 H. pylori
3	15	100.0	360	4	Aaf88153 H. pylori
4	15	100.0	360	4	Aaf88096 H. pylori
5	15	100.0	36568	6	Abk50980 Human sol
6	15	100.0	53809	10	Adc86800 Human GPC
7	15	100.0	128361	11	Adn95229 Human BEC
8	14	93.3	185	3	Aax82871 Human dys
9	14	93.3	185	3	Aaa36780 Human dys
10	14	93.3	247	6	Adv89025 Human col
11	14	93.3	498	6	Abn96062 Gene #256
12	14	93.3	537	3	Aaa79498 Eucalyptu
13	14	93.3	848	3	Aaa79497 Eucalyptu
14	14	93.3	1040	4	Aak89212 Human dig
15	14	93.3	1040	4	Aak89213 Human dig
16	14	93.3	2447	3	Aaa79719 Eucalyptu
17	14	93.3	2757	13	Adq84612 Human tum
18	14	93.3	2758	10	Ada53027 Human cod
19	14	93.3	2843	11	Adm02813 Human cDN
20	14	93.3	4374	4	Aak52302 Human pol

C 21	14	93.3	4455	4	AAK52301 Human pol
C 22	14	93.3	4580	3	Aac77082 Human ORF
C 23	14	93.3	5092	12	Adj27336 Human NPC
C 24	14	93.3	5180	4	AAK82269 Human imm
C 25	14	93.3	21501	10	ACF79823 Human SOS
C 26	14	93.3	52640	12	Adq97220 Mouse can
C 27	14	93.3	80450	3	Aaf22295 BAC conta
C 28	14	93.3	94752	10	Adf11646 Human chr
C 29	14	93.3	110000	8	AdS52230
C 30	14	93.3	130320	10	Adf11613 Human scl
C 31	14	93.3	154681	10	Adl13861 Osteoar
C 32	14	93.3	160755	4	AAH88704 Human DNA
C 33	14	93.3	204803	12	Adq97348 Mouse can
C 34	14	93.3	213040	11	ACM44418 Human gen
C 35	14	93.3	214019	10	Adl13809 Osteoar
C 36	14	93.3	252907	13	ABD32694 Human can
C 37	13.4	89.3	24	2	AAQ43233 B-810 V r
C 38	13.4	89.3	45	10	ABZ68657 Primer us
C 39	13.4	89.3	53	12	ADQ31301 Mutagenic
C 40	13.4	89.3	54	10	ADT25193 Humanised
C 41	13.4	89.3	54	12	ADF94357 Anti VIA-
C 42	13.4	89.3	56	12	ADQ31298 Mutagenic
C 43	13.4	89.3	66	10	ABZ68633 Primer fo
C 44	13.4	89.3	123	3	AAC25363 Human sec
C 45	13.4	89.3	279	9	ACD27478 Mouse DNA

## ALIGNMENTS

### RESULT 1

AAF88108

ID AAF88108 standard; DNA; 15 BP.

XX

AC AAF88108;

XX

DT 17-JUL-2001 (first entry)

XX

DE H. pylori catalase derived antibody HP25/6m/1B5 heavy chain CDR1 DNA.

XX

KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;

KM acid-resistant microorganism; complementarity determining region; CDR;

KW feces; heavy chain; light chain; ds.

XX

OS Unidentified.

XX

PN WC200127612-A2.

XX

PD 19-APR-2001.

XX

PF 12-OCT-2000; 2000WO-EP010057.

XX

PR 12-OCT-1999; 99EP-00120351.

PR

PR 16-MAR-2000; 2000EP-00105592.

PR

PR 31-MAR-2000; 2000EP-00107028.

XX

XX 10-MAY-2000; 2000EP-00110110.

XX

PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX

PI Reiter C, Cullmann G, Lakner M, Truse A, Dehnert S, Schwartz G;

XX

DR WFI; 2001-282086/29.

DR

P-PSDB; AAB86081.

XX

XX Detecting infections by acid-resistant microorganisms, particularly for

PT diagnosing Helicobacter pylori, comprises immunochromatographic detection

PT of antigen in feces.

XX

PS Claim 26; Page 25; 90pp; German.

XX

CC This invention describes a novel method for detecting infection by an

CC acid-resistant microorganism (A), in a mammal, using

CC immunochromatography. The method is used to diagnose infection by an acid

CC -resistant microorganism (A), in a mammal, such as *Helicobacter*,  
 CC *Campylobacter* or *Mycobacterium*, particularly *H. pylori* (most preferred),  
 CC *H. hepatica*, *C. jejuni* and *M. tuberculosis*. The method is rapid, simple,  
 CC inexpensive and non-invasive, and may indicate the stage of infection. A  
 CC test strip used in the method may include a filter to eliminate particles  
 CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence encodes a  
 CC complementarity determining region (CDR) from an antibody raised against  
 CC the *H. pylori* catalase or beta-urease antigen which is used to illustrate  
 CC the method of the invention

XX SQ Sequence 15 BP; 4 A; 5 C; 3 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 15; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCAC 15  
 |||||  
 Db 1 GACACCTATGTGCAC 15

## RESULT 2

AAF88051  
 ID AAF88051 standard; DNA; 15 BP.

XX AAF88051;

DT 17-JUL-2001 (first entry)

XX H. pylori catalase derived antibody HP25/6m/1B5 heavy chain CDR1 DNA.

XX Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
 KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
 KW antibacterial; complementarity determining region; ds.

XX Unidentified.

XX WO200127613-A2.

XX 19-APR-2001.

XX 12-OCT-2000; 2000WO-EP010058.

XX 12-OCT-1999; 99EP-00120351.

PR 16-MAR-2000; 2000EP-00105592.

PR 31-MAR-2000; 2000EP-00107028.

PR 10-MAY-2000; 2000EP-00110110.

XX (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX Reiter C, Cullmann G, Heppner P, Ringels A, Mueller H, Haindl E;

XX WPI; 2001-282087/29.

XX P-PSDB; AAB86049.

XX Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing *Helicobacter pylori*, comprises an immunoassay on a fecal  
 PT sample.

XX Claim 20; Page 16; 89pp; German.

XX This invention describes a novel method for detecting, in a mammal,  
 CC infection by an acid-resistant microorganism (A) which comprises reacting  
 CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
 CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
 CC complex is formed with Ag, and the formation of a complex detected. R are  
 CC specific for an Ag which, after passage through the intestines, at least  
 CC in some mammals, retains a native (or corresponding) structure against  
 CC which the mammal produces antibodies (when immunized or infected with  
 CC (A), or its extracts, lysates or derived proteins (or fragments) or

CC synthetic peptides). The products of the invention have antibacterial  
 CC activity. The method is used to diagnose infection by *Helicobacter*,  
 CC *Campylobacter* or *Mycobacterium*, particularly *H. pylori* (most preferred),  
 CC *H. hepatica*, *C. jejuni* and *M. tuberculosis*, and also to monitor the  
 CC progress of treatment. Receptors, particularly antibodies, directed  
 CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an  
 CC infection. This sequence encodes a complementarity determining region  
 CC (CDR) from an antibody generated against a *Helicobacter pylori* antigen  
 CC (catalase or beta-urease) which is used to illustrate the method of the  
 CC invention

XX SQ Sequence 15 BP; 4 A; 5 C; 3 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 15; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCAC 15  
 |||||  
 Db 1 GACACCTATGTGCAC 15

## RESULT 3

AAF88153

ID AAF88153 standard; DNA; 360 BP.

XX AAF88153;

DT 17-JUL-2001 (first entry)

XX H. pylori catalase antibody HP25/6m/1B50 V region heavy chain DNA.

XX Catalase; beta-urease; antibody; antigen; detection; infection; epitope;  
 KW acid-resistant microorganism; complementarity determining region; CDR;  
 KW feces; heavy chain; light chain; ds.

XX Unidentified.

XX WO200127612-A2.

XX 19-APR-2001.

XX 12-OCT-2000; 2000WO-EP010057.

XX 12-OCT-1999; 99EP-00120351.

PR 16-MAR-2000; 2000EP-00105592.

PR 31-MAR-2000; 2000EP-00107028.

PR 10-MAY-2000; 2000EP-00110110.

XX (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;

XX WPI; 2001-282086/29.

XX P-PSDB; AAB86107.

XX Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing *Helicobacter pylori*, comprises immunochromatographic detection  
 PT of antigen in feces.

XX Claim 32; Fig 3; 90pp; German.

XX This invention describes a novel method for detecting infection by an  
 CC acid-resistant microorganism (A), in a mammal, using  
 CC immunochromatography. The method is used to diagnose infection by an acid  
 CC -resistant microorganism (A), in a mammal, such as *Helicobacter*,  
 CC *Campylobacter* or *Mycobacterium*, particularly *H. pylori* (most preferred),  
 CC *H. hepatica*, *C. jejuni* and *M. tuberculosis*. The method is rapid, simple,  
 CC inexpensive and non-invasive, and may indicate the stage of infection. A

CC test strip used in the method may include a filter to eliminate particles  
 CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence encodes a  
 CC Helicobacter pylori catalase derived antibody V-region heavy chain  
 CC fragment used to illustrate the method of the invention  
 XX  
 XX Sequence 360 BP; 87 A; 95 C; 90 G; 88 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 15; DB 4; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACACCTATGTGCAC 15  
 DB 91 GACACCTATGTGCAC 105  
 RESULT 4  
 AAF88096  
 ID AAF88096 standard; DNA; 360 BP.  
 XX AC AAF88096;  
 XX DT 17-JUL-2001 (first entry)  
 XX DE H. pylori catalase antibody HP25/6m/1B5 V-region heavy chain DNA.  
 XX KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
 KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
 KW antibacterial; complementarity determining region; V-region; db.  
 XX OS Unidentified.  
 XX PN WO200127613-A2.  
 XX PD 19-APR-2001.  
 XX PF 12-OCT-2000; 2000WO-EP010058.  
 XX PR 12-OCT-1999; 99EP-00120351.  
 XX PR 16-MAR-2000; 2000EP-00105592.  
 XX PR 31-MAR-2000; 2000EP-00107028.  
 XX PR 10-MAY-2000; 2000EP-00110110.  
 XX PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
 XX PI Reiter C, Cullmann G, Heppner P, Ringeis A, Mueller H, Haindl E;  
 XX DR WPI; 2001-282087/29.  
 XX DR P-PSDB; AAB86069.  
 XX PT Detecting infections by acid-resistant microorganisms, particularly for  
 XX diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
 XX sample.  
 XX PS Claim 28; Fig 3; 89pp; German.  
 XX CC This invention describes a novel method for detecting, in a mammal,  
 CC infection by an acid-resistant microorganism (A) which comprises reacting  
 CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
 CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
 CC complex is formed with Ag, and the formation of a complex detected. R are  
 CC specific for an Ag which, after passage through the intestines, at least  
 CC in some mammals, retains a native (or corresponding) structure against  
 CC which the mammal produces antibodies (when immunized or infected with  
 CC (A), or its extracts, lysates or derived proteins (or fragments) or  
 CC synthetic peptides). The products of the invention have antibacterial  
 CC activity. The method is used to diagnose infection by Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the  
 CC progress of treatment. Receptors, particularly antibodies, directed

CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an  
 CC infection. This sequence encodes a Helicobacter pylori anti-catalase  
 CC derived antibody HP25/6m/1B5 V-region heavy chain fragment which is  
 CC described in the method of the invention  
 XX  
 XX Sequence 360 BP; 87 A; 95 C; 90 G; 88 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 15; DB 4; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACACCTATGTGCAC 15  
 DB 91 GACACCTATGTGCAC 105  
 RESULT 5  
 ABK50980  
 ID ABK50980 standard; DNA; 36568 BP.  
 XX AC ABK50980;  
 XX DT 24-SEP-2002 (first entry)  
 XX DE Human solute carrier family 1 gene, SLC1A4.  
 XX KW Solute carrier family 1; SLC1A4; haplotyping; human; cancer; gene;  
 KW glutamate/neutral amino acid transporter; neurological disease; ds;  
 KW amino acid transporter disorder; single nucleotide polymorphism; SNP.  
 XX OS Homo sapiens.  
 XX PH Key Location/Qualifiers  
 FT intron 1..4017  
 FT /\*tag= a  
 FT /number= 1  
 FT replace(3627, T)  
 FT /\*tag= b  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT replace(3639, A)  
 FT /\*tag= c  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT replace(3738, G)  
 FT /\*tag= d  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT replace(3879, C)  
 FT /\*tag= e  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT replace(3887, G)  
 FT /\*tag= f  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT replace(3888, C)  
 FT /\*tag= g  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT replace(3967, C)  
 FT /\*tag= h  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT replace(3983, T)  
 FT /\*tag= i  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT 4018..35520  
 FT /\*tag= k  
 FT /product= "Solute carrier family 1 protein, SLC1A4"  
 FT 4018..4544  
 FT /\*tag= j  
 FT /number= 1  
 FT replace(4038, A)  
 FT /\*tag= l  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation

```
FT variation replace(4126, C)
FT /*tag= m
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(4251, G)
FT /*tag= n
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(4329, C)
FT /*tag= o
FT /standard_name= "Single nucleotide polymorphism"
FT intron 4545..15821
FT /*tag= p
FT /number= 2
FT variation replace(15699, T)
FT /*tag= q
FT /standard_name= "Single nucleotide polymorphism"
FT exon 15822..15864
FT /*tag= x
FT /number= 2
FT intron 15865..18326
FT /*tag= s
FT /number= 3
FT variation replace(15946, A)
FT /*tag= t
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(18199, A)
FT /*tag= u
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(18262, C)
FT /*tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT exon 18327..18389
FT /*tag= w
FT /number= 3
FT intron 18690..24970
FT /*tag= x
FT /number= 4
FT exon 24971..25137
FT /*tag= y
FT /number= 4
FT intron 25138..30813
FT /*tag= z
FT /number= 5
FT exon 30814..31047
FT /*tag= aa
FT /number= 5
FT intron 31048..32444
FT /*tag= ab
FT /number= 6
FT exon 32445..32639
FT /*tag= ac
FT /number= 6
FT variation replace(32480, C)
FT /*tag= ad
FT /standard_name= "Single nucleotide polymorphism"
FT intron 32640..32918
FT /*tag= ae
FT /number= 7
FT variation replace(32831, T)
FT /*tag= af
FT /standard_name= "Single nucleotide polymorphism"
FT exon 32919..33053
FT /*tag= ag
FT /number= 7
FT intron 33054..335285
FT /*tag= ah
FT /number= 8
FT variation replace(33129, A)
FT /*tag= ai
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(33157, T)
FT /*tag= aj
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(35243, A)
```

```
FT /*tag= ak
FT /standard_name= "Single nucleotide polymorphism"
FT exon 35286..35520
FT /*tag= al
FT /number= 8
FT variation replace(35511, A)
FT /*tag= am
FT /standard_name= "Single nucleotide polymorphism"
FT intron 35521..36568
FT /*tag= an
FT /number= 9
FT XX
FT PN W0200244198-A2.
FT XX
FT PD 06-JUN-2002.
FT XX
FT PF 29-NOV-2001; 2001WO-US044781.
FT PR 30-NOV-2000; 2000US-0250254P.
FT XX (GENA-) GENAISSANCE PHARM INC.
FT XX
FT PI Bieglecki KM, Kazemi A, Russo DP, Sausker EA;
FT XX
FT DR WPI; 2002-519580/55.
FT DR P-PSDB; AAU80097.
FT XX
FT PT Novel genetic variants of Solute Carrier Family 1 (Glutamate/Neutral
FT Amino Acid Transporter), Member 4 isogenes, for improving efficiency and
FT reliability in drug development for treating cancers.
FT XX
FT PS Claim 20; Fig 1; 139pp; English.
FT XX
FT CC The invention relates to an isolated polynucleotide (I) comprising a
FT first nucleotide sequence which comprises solute carrier family 1
FT (glutamate/neutral amino acid transporter), member 4 (SLC1A4) isogenes
FT (II) and an isolated polypeptide (III) comprising an amino acid sequence
FT which is a polymorphic variant of a reference sequence for SLC1A4
FT protein. Also described are methods for: (1) haplotyping or genotyping
FT SLC1A4 gene of an individual; (2) predicting a haplotype pair for SLC1A4
FT gene of an individual; (3) identifying an association between a trait and
FT at least one haplotype or haplotype pair of SLC1A4 gene. (III) is useful
FT in screening for drugs targeting (III) that are useful for treating
FT cancer, neurological diseases and amino acid transporter disorders. The
FT methods are useful for improving the efficiency and reliability of
FT several steps in the discovery and development of drugs for treating
FT diseases associated with SLC1A4 activity. The haplotyping method is also
FT used by the pharmaceutical research scientist to validate SLC1A4 as a
FT candidate target for treating a specific condition or disease predicted
FT to be associated with SLC1A4 activity, e.g. cancer, neurological diseases
FT and amino acid transporter disorders, and in the design of clinical
FT trials for treating a specific condition of disease associated with
FT SLC1A4 activity. The methods are also useful for screening compounds
FT targeting SLC1A4. Anti-SLC1A4 antibody is useful in diagnostic,
FT prognostic and therapeutic methods. The present sequence represents the
FT genomic DNA sequence of SLC1A4
FT XX
FT SQ Sequence 36568 BP; 10127 A; 8102 C; 8482 G; 9827 T; 0 U; 30 Other;
FT
FT Query Match 100.0%; Score 15; DB 6; Length 36568;
FT Best Local Similarity 100.0%; Pred. No. 1.5e+02;
FT Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
FT QY 1 GACACCTATGTGCAC 15
FT |||||
FT Db 17692 GACACCTATGTGCAC 17706
FT
FT RESULT 6
FT ADC86800
FT ID ADC86800 standard; DNA; 53809 BP.
FT XX
FT AC ADC86800;
```



```

XX 01-JAN-2004 (first entry)
XX Human GPCR gene SEQ ID NO:1253.
XX ds; gene; human; GPCR;
XX guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX Homo sapiens.
XX EP1270724-A2.
XX 02-JAN-2003.
XX 18-JUN-2002; 2002EP-00013517.
XX 18-JUN-2001; 2001JP-00246789.
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX Suwa M, Asai K, Akiyama Y, Aburatani H;
XX WPI; 2003-315783/31.
XX P-PSDB; ADC86801.
XX New polynucleotide, useful for preparing a composition for treating a
XX patient in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor.
XX Claim 1; SEQ ID NO 1253; 28pp; English.
XX The invention relates to a novel polynucleotide encoding a guanosine
XX triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
XX the invention may have a use in gene therapy. The polynucleotide and
XX polypeptide are useful for preparing a composition for treating a patient
XX in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor. The
XX polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
XX invention.
XX Sequence 53809 BP; 9635 A; 17151 C; 15840 G; 11183 T; 0 U; 0 Other;
Query Match 100.0%; Score 15; DB 10; Length 53809;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACACCTATGTGCAC 15
DB 4195 GACACCTATGTGCAC 4209
|||||
RESULT 7
ADN95229
ID ADN95229 standard; DNA; 128361 BP.
XX AC ADN95229;
XX 01-JUL-2004 (first entry)
XX Human BEC/LEC-related gene sequence SeqID151.
XX growth; differentiation; blood endothelial cell; BEC;
XX lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
XX lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;
XX vasotrophic; antiinflammatory; gene therapy; endothelial cell disorder;
XX inflammatory disease; cancer metastasis; lymphatic system; gene; ds;
XX human.
XX Homo sapiens.
XX OS
XX WO2003080640-A1.
XX PN
XX

```

```

PD 02-OCT-2003.
XX 07-MAR-2003; 2003WO-US006900.
XX PF
XX 07-MAR-2002; 2002US-0363019P.
XX PR
XX (LUDW-) LUDWIG INST CANCER RES.
XX PA (LICN) LICENTIA LTD.
XX PI
XX Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;
XX WPI; 2003-876899/81.
XX DR P-PSDB; ADN95228.
XX DR
XX Example 1; SEQ ID NO 151; 176pp; English.
XX PS
XX This invention relates to a method of differentially modulating the
XX growth or differentiation of blood endothelial cells (BEC) or lymphatic
XX endothelial cells (LEC) comprises contacting endothelial cells with a
XX composition comprising an agent that differentially modulates blood or
XX lymphatic endothelial cells. Treating hereditary lymphoedema comprises
XX identifying a human subject with lymphoedema and with a mutation in at
XX least one allele of a gene encoding a LEC protein, where the mutation
XX correlates with lymphoedema in human subjects, and with the proviso that
XX the LEC protein is not VEGFR-3; and administering to the subject a
XX composition comprising a lymphatic growth agent selected from VEGF-C or
XX VEGF-D polypeptides and polynucleotides. The invention may be useful for
XX the development of compounds with an antiangiogenic, cytostatic,
XX vasotrophic or antiinflammatory activity or for gene therapy. The method
XX is useful in modulating the growth or differentiation of blood
XX endothelial cells or lymphatic endothelial cells, in treating hereditary
XX lymphoedema, in screening for an endothelial cell disorder or
XX predisposition to the disorder or in monitoring the efficacy or toxicity
XX of a drug on endothelial cells. The agent is useful in manufacturing a
XX medicament for the differential modulation of blood vessel endothelial
XX cell or lymphatic vessel endothelial cell growth or differentiation. The
XX lymphatic growth agent may also be used in manufacturing a medicament for
XX the treatment of hereditary lymphoedema resulting from a mutation in a
XX LEC gene or of other diseases involving the lymphatic vessels, such as
XX various inflammatory diseases and cancer metastasis via the lymphatic
XX system. The present sequence is that of a human LEC/BEC differentially
XX expressed gene which is related to the method of the invention. Note: This
XX sequence does not appear in the specification but was obtained by the
XX indexer using the source data given in table 14 of the specification.
XX SQ Sequence 128361 BP; 26995 A; 36536 C; 34778 G; 30052 T; 0 U; 0 Other;
Query Match 100.0%; Score 15; DB 11; Length 128361;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACACCTATGTGCAC 15
DB 37247 GACACCTATGTGCAC 37261
|||||
RESULT 8
AAx82871/c
ID AAx82871 standard; DNA; 185 BP.
XX AC AAx82871;
XX 30-JUN-2000 (first entry)
XX DT
XX Human dysferlin DNA #19.
XX DE
XX Dysferlin; anti-dystrophic; gene therapy; muscular dystrophy; human;
XX skeletal muscle cell; hereditary; Miyoshi myopathy; diagnosis;
XX limb girdle muscular dystrophy-2b; brain-specific; ss.
XX KW
XX Homo sapiens.
XX OS
XX WO200011157-A1.
XX PN

```

```

XX PD 02-MAR-2000.
XX XX
XX PF 25-AUG-1999; 99WO-US019395.
XX XX
XX PR 25-AUG-1998; 98US-0097927P.
XX XX
XX PA (GEHO ) GEN HOSPITAL CORP.
XX PI Brown RH, Liu J, Aoki M, Ho MF, Mateuda-Asada C;
XX DR WPI; 2000-237646/20.
XX XX
XX PT Novel dyserferlin genes and related proteins useful for diagnosis, risk
XX PT identification and treatment of hereditary muscular dystrophies and other
XX PT dyserferlin related disorders.
XX XX
XX PS Claim 11; Page 110; 146pp; English.
XX CC
XX CC This invention describes a novel human dyserferlin nucleic acid (I) and its
XX CC encoding protein (II), which has anti-dystrophic activity and can be used
XX CC for gene therapy. Introduction of (I), a vector comprising (I) or
XX CC dyserferlin into a cell of a mammal can be used to decrease the symptoms of
XX CC muscular dystrophy. The dyserferlin gene is normally expressed in skeletal
XX CC muscle cells and is selectively mutated in several families with the
XX CC hereditary muscular dystrophies, e.g. Miyoshi myopathy and limb girdle
XX CC muscular dystrophy-2B. The primers and oligonucleotides derived from (I)
XX CC can be used in diagnosis of or risk identification for dyserferlin-related
XX CC disorders in patients, fetus, or pre-embryos. Expression of brain-
XX CC specific dyserferlin may be important as a marker for normal neural
XX CC development. Dyserferlin DNA or subgenomic coding sequences can be used for
XX CC therapy of the hereditary muscular dystrophies. This sequence represents
XX CC a fragment of the human dyserferlin gene described in the method of the
XX CC invention
XX XX
XX SQ Sequence 185 BP; 36 A; 47 C; 44 G; 58 T; 0 U; 0 Other;

Query Match 93.3%; Score 14; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14
Db |||||
176 GACACCTATGTGCA 163

RESULT 9
AAA36780/C
ID AAA36780 standard; DNA; 185 BP.
AC AAA36780;
XX
XX 03-AUG-2000 (first entry)
XX
XX Human dyserferlin related nucleotide sequence SEQ ID NO:38.
XX
XX Human; dyserferlin; mutant; identification; Chromosome 2p12-14; detection;
XX KW muscular dystrophy; diagnosis; hereditary muscular dystrophy;
XX KW miyoshi myopathy; limb girdle muscular dystrophy; ds.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200011016-A1.
XX XX
XX PD 02-MAR-2000.
XX XX
XX PF 25-AUG-1999; 99WO-US019394.
XX XX
XX PR 25-AUG-1998; 98US-0097930P.
XX XX
XX PR (GEHO ) GEN HOSPITAL CORP.
XX PA (UYPI-) UNIV PITTSBURGH.
XX XX

Brown RH, Liu J, Hoffman E, Chou F;
WPI; 2000-246531/21.
Dysferlin polynucleotide, its mutant form useful for diagnosis and
treatment of hereditary muscular dystrophies e.g. miyoshi myopathy and
limb girdle muscular dystrophy.
Disclosure; Page 100; 136pp; English.
The present invention describes an isolated dysferlin DNA of 20-25
nucleotides in length, comprising a nucleotide sequence specifically
selected from nucleotides 911-913, 929-948, 1019-1038, 1392-1411, 1424-
1443, 1484-1503, 1499-1518, 1543-1565, 1715-1734, 1714-1759, 2241-2260,
2864-2883, 2978-2997, 3057-3076, 3198-3217, 3252-3271, 4356-4375, 4665-
4684, 5015-5034, 5610-5629, 5726-5735, 6035-6054, 6179-6198, 6243-6263
and 6529-6548 of the human dysferlin nucleotide sequence given in
AAA36744. Dyserferlin nucleotide sequences containing specific mutations
can be used for diagnosing a patient, a fetus or a pre-embryo at risk of
developing a dyserferlin associated disorder by detecting mutations in the
dysferlin gene in biological samples from patients. Alternatively, the
biological sample containing genomic DNA can be incubated with a
restriction enzyme, preferably BspII, BspI286I, RsaI, HhaI, HaeIII,
BspI286I, NlaIV, NciII, BclI, AclI, PstI, SfiI, PfuI, AclI, AclI, AclI,
Tsp509I, SalI, HincII, TagI, HinfI, TfiI, SfiI or FokI and the presence
or absence of a restriction enzyme site in the sample is detected as an
indication of the presence or absence of a particular mutation in the
sample. Dyserferlin polynucleotides are useful for treating hereditary
muscular dystrophies such as miyoshi myopathy (MM) and limb girdle
muscular dystrophy-2B (LGM2B-2B). MM and LGM2B-2B map to the human
chromosome 2p12-14 region between the genetic markers D2S292 and D2S286.
The present sequence represents a nucleotide sequence given in the
sequence listing of the present specification, but which does not appear
to be mentioned further
Sequence 185 BP; 36 A; 47 C; 44 G; 58 T; 0 U; 0 Other;

Query Match 93.3%; Score 14; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14
Db |||||
176 GACACCTATGTGCA 163

RESULT 10
ABV89025
ID ABV89025 standard; cDNA; 247 BP.
AC ABV89025;
XX
XX 13-DEC-2002 (first entry)
XX
XX Human colon cancer related cDNA SEQ ID NO 2336.
XX
XX Human; colon; cancer; cytostatic; tumour; gene therapy; vaccine; gene;
XX KW ss.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200258534-A2.
XX XX
XX PD 01-AUG-2002.
XX XX
XX PF 16-NOV-2001; 2001WO-US043704.
XX XX
XX PR 20-NOV-2000; 2000US-0252222P.
XX PR 06-FEB-2001; 2001US-0267011P.
XX PR 28-MAR-2001; 2001US-0279670P.
XX PR 10-JUL-2001; 2001US-0304037P.
XX XX
XX PA (CORI-) CORIXA CORP.

```

XX Stolk JA, Xu J, Chenault RA, Meagher MJ, Secrist H, King GE;  
XX WPI; 2002-608400/65.  
XX New isolated tumor colon polynucleotide and polypeptide, useful for the  
XX diagnosis, prevention and/or treatment of cancer, in particular colon  
XX cancer.  
XX Claim 1; SEQ ID NO 2336; 266pp + Sequence Listing; English.  
XX The invention relates to a human colon tumour expressed polynucleotide  
XX (I) encoding a polypeptide (II, ABP67991-ABP67996) comprising: (i) any of  
XX 2600 fully defined nucleotide sequences (ABV8669-ABV89289); (ii)  
XX complements of (i); (iii) at least 20 contiguous residues of (i); (iv)  
XX sequences that hybridize to (i), under moderately stringent conditions;  
XX (v) sequences having at least 75% or 90% identity to (i); or (vi)  
XX degenerate variants of (i). The compositions and methods of the present  
XX invention are useful for the diagnosis, prevention and/or treatment of  
XX cancer, particularly colon cancer. (I) can be used in gene therapy and  
XX (i) and (ii) are useful in pharmaceutical compositions such as vaccines.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Query Match 93.3%; Score 14; DB 6; Length 247;  
XX Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1 GACACCTATGTGCA 14  
XX 117 GACACCTATGTGCA 130  
XX  
XX RESULT 11  
XX ABN96062  
XX ID ABN96062 standard; DNA; 498 BP.  
XX AC ABN96062;  
XX  
XX 13-AUG-2002 (first entry)  
XX Gene #2560 used to diagnose liver cancer.  
XX  
XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
XX metastatic liver tumour; cytostatic; expression profile; disease state;  
XX disease progression; drug toxicity; drug efficacy; drug metabolism.  
XX  
XX Homo sapiens.  
XX  
XX WO200229103-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 02-OCT-2001; 2001WO-US030589.  
XX  
XX 02-OCT-2000; 2000US-0237054P.  
XX  
XX (GENE-) GENE LOGIC INC.  
XX  
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
XX WPI; 2002-426119/45.  
XX  
XX Diagnosing and detecting the progression of liver cancer, hepatocellular  
XX carcinoma or metastatic liver tumor in a patient, involves detecting the  
XX level of expression of two or more genes in a liver tissue sample.  
XX  
XX Claim 1; SEQ ID NO 2560; 298pp; English.  
XX  
XX The invention relates to a novel method for diagnosing and detecting the

CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
CC tumour in a patient, and differentiating metastatic liver cancer from  
CC hepatocellular carcinoma in a patient, involving detecting the level of  
CC expression of two or more genes represented in ABN93503-ABN97455 in a  
CC tissue sample. The method of the invention has hepatotropic, and  
CC cytostatic activity. The method is useful for diagnosing and detecting  
CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
CC liver carcinoma in a patient. The method is useful for identifying  
CC expression profiles which serve as useful diagnostic markers as well as  
CC markers that can be used to monitor disease states, disease progression,  
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Query Match 93.3%; Score 14; DB 6; Length 498;  
XX Best Local Similarity 100.0%; Pred. No. 4e+02;  
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 2 ACACCTATGTGCA 15  
XX 36 ACACCTATGTGCA 49  
XX  
XX RESULT 12  
XX AAA79498/c  
XX ID AAA79498 standard; cDNA; 537 BP.  
XX AC AAA79498;  
XX  
XX 27-NOV-2000 (first entry)  
XX  
XX Eucalyptus grandis cell signalling involved polynucleotide SEQ ID NO:299.  
XX  
XX Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
XX plant cell signalling; modulation; transgenic plant; pathogen; growth;  
XX environmental change; development; cell proliferation; differentiation;  
XX elongation; survival; disease resistance; nutrient metabolism; ss.  
XX  
XX Eucalyptus grandis.  
XX  
XX WO200042171-A1.  
XX  
XX 20-JUL-2000.  
XX  
XX 11-JAN-2000; 2000WO-US000724.  
XX  
XX 12-JAN-1999; 99US-00228986.  
XX  
XX 01-NOV-1999; 99US-0162866P.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX Strabala TJ, Nieuwenhuizen NJ;  
XX WPI; 2000-476052/41.  
XX  
XX Isolated polynucleotide encoding a polypeptide involved in cell signaling  
XX used for generating transgenic plants with modified responses to external  
XX signals.  
XX  
XX Claim 1; Page 155; 527pp; English.  
XX  
XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide  
XX and protein sequences isolated from eucalyptus (Eucalyptus grandis) or  
XX pine (Pinus radiata) also known as Monterey pine. The protein sequences  
XX are involved in cell signalling. The polynucleotide and protein sequences  
XX can be used to modify the response of plant cells to external signals  
XX e.g. environmental changes or pathogens during the growth and development  
XX of a plant. They can be used to modify cell proliferation,  
XX differentiation, elongation and survival, resistance to disease and  
XX nutrient metabolism. Examples of modifications which can be produced are

CC altered fruit ripening and senescence of leaves and flowers e.g. to delay  
CC senescence and prolong the life of cut flowers or enhance senescence of  
CC reproductive organs to engineer sterile plants. Other modifications can  
CC be used to delay senescence in selected cell types or organs providing  
CC fruit and vegetables which have a longer shelf life between harvest and  
CC consumption, or to decrease branching frequency in forest tree species  
CC giving long stretches of valuable knot-free clear wood which can be used  
CC in solid timber furniture and veneers  
XX  
SQ Sequence 537 BP; 136 A; 136 C; 151 G; 114 T; 0 U; 0 Other;  
Query Match 93.3%; Score 14; DB 3; Length 537;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 ACACCTATGTGCAC 15  
Db 156 ACACCTATGTGCAC 143  
|||||  
RESULT 13  
AAA79497/c  
ID AAA79497 standard; cDNA; 848 BP.  
XX  
AC AAA79497;  
XX  
DT 27-NOV-2000 (first entry)  
XX  
DE Eucalyptus grandis cell signalling involved polynucleotide SEQ ID NO:298.  
XX  
KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;  
KW environmental change; development; cell proliferation; differentiation;  
KW elongation; survival; disease resistance; nutrient metabolism; ss.  
XX  
OS Eucalyptus grandis.  
XX  
PN WO200042171-A1.  
XX  
PD 20-JUL-2000.  
XX  
PF 11-JAN-2000; 2000WO-US000724.  
XX  
PR 12-JAN-1999; 99US-00228986.  
PR 01-NOV-1999; 99US-0162866P.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
PI Strabala TJ, Nieuwenhuizen NJ;  
XX  
DR WPI; 2000-476052/41.  
XX  
PT Isolated polynucleotide encoding a polypeptide involved in cell signaling  
PT used for generating transgenic plants with modified responses to external  
PT signals.  
XX  
XX  
XX Claim 1; Page 155; 527pp; English.  
XX  
CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide  
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or  
CC pine (Pinus radiata also known as Monterey pine). The protein sequences  
CC are involved in cell signalling. The polynucleotide and protein sequences  
CC can be used to modify the response of plant cells to external signals  
CC e.g. environmental changes or pathogens during the growth and development  
CC of a plant. They can be used to modify cell proliferation,  
CC differentiation, elongation and survival, resistance to disease and  
CC nutrient metabolism. Examples of modifications which can be produced are  
CC altered fruit ripening and senescence of leaves and flowers e.g. to delay  
CC senescence and prolong the life of cut flowers or enhance senescence of  
CC reproductive organs to engineer sterile plants. Other modifications can  
CC be used to delay senescence in selected cell types or organs providing  
CC fruit and vegetables which have a longer shelf life between harvest and  
CC consumption, or to decrease branching frequency in forest tree species

CC giving long stretches of valuable knot-free clear wood which can be used  
CC in solid timber furniture and veneers  
XX  
SQ Sequence 848 BP; 209 A; 218 C; 230 G; 191 T; 0 U; 0 Other;  
Query Match 93.3%; Score 14; DB 3; Length 848;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 ACACCTATGTGCAC 15  
Db 142 ACACCTATGTGCAC 129  
|||||  
RESULT 14  
AAK89212/c  
ID AAK89212 standard; DNA; 1040 BP.  
XX  
AC AAK89212;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human digestive system antigen genomic sequence SEQ ID NO: 2788.  
XX  
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;  
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;  
KW digestive system disorder; Meckel's diverticulum; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155314-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001324.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216847P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 30-AUG-2000; 2000US-0227009P.  
PR 01-SEP-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228927P.

PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 08-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX  
XX WPI; 2001-502630/55.  
XX  
PT Polynucleotides encoding digestive system antigens, useful for  
PT diagnosing, treating, preventing and/or prognosing disorders of the  
PT digestive system, particularly cancer and cancer metastases.  
XX  
PS Disclosure; SEQ ID NO 2788; 986pp; English.  
XX  
CC The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic  
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a genomic DNA fragment  
CC encoding a digestive system antigen of the invention  
XX  
SQ Sequence 1040 BP; 228 A; 308 C; 262 G; 242 T; 0 U; 0 Other;

Query Match 93.3%; Score 14; DB 4; Length 1040;  
Best Local Similarity 100.0%; Pred. NO. 4.3e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACCTATGTGCAC 15  
|||||  
Db 1016 ACACCTATGTGCAC 1003

RESULT 15  
AAK89213/c  
ID AAK89213 standard; DNA; 1040 BP.  
XX  
AC AAK89213;  
XX

05-NOV-2001 (first entry)

Human digestive system antigen genomic sequence SEQ ID NO: 2789.

Human; digestive system antigen; gene therapy; cancer; appendicitis;  
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;  
KW digestive system disorder; Meckel's diverticulum; ds.  
XX  
OS Homo sapiens.

XX  
PN WO200155314-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US0011324.  
XX  
PP  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 14-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 21-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 25-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 26-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 08-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2001US-0254097P.  
PR  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX

PI Rosen CA, Barash SC, Ruben SM;  
DR WPI; 2001-502630/55.  
XX  
PT Polynucleotides encoding digestive system antigens, useful for  
PT diagnosing, treating, preventing and/or prognosing disorders of the  
PT digestive system, particularly cancer and cancer metastases.  
XX  
XX Disclosure; SEQ ID NO 2789; 986pp; English.  
PS  
XX The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic  
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a genomic DNA fragment  
CC encoding a digestive system antigen of the invention  
XX  
SQ Sequence 1040 BP; 228 A; 309 C; 262 G; 241 T; 0 U; 0 Other;  
Query Match 93.3%; Score 14; DB 4; Length 1040;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 ACACCTATGTGCAC 15  
DB 1016 ACACCTATGTGCAC 1003  
RESULT 16  
AAA79719/C  
ID AAA79719 standard; cDNA; 2447 BP.  
XX  
AC AAA79719;  
XX  
DT 27-NOV-2000 (first entry)  
XX  
DE Eucalyptus grandis cell signalling involved polynucleotide SEQ ID NO:871.  
XX  
KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;  
KW environmental change; development; cell proliferation; differentiation;  
KW elongation; survival; disease resistance; nutrient metabolism; ss.  
XX  
OS Eucalyptus grandis.  
XX  
PN WO200042171-A1.  
XX  
PD 20-JUL-2000.  
XX  
PF 11-JAN-2000; 2000WO-US000724.  
XX  
PR 12-JAN-1999; 99US-00228986.  
PR 01-NOV-1999; 99US-0162866P.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
PI Strabala TJ, Nieuwenhuizen NJ;  
XX  
XX WPI; 2000-476052/41.  
XX  
PT Isolated polynucleotide encoding a polypeptide involved in cell signaling  
PT used for generating transgenic plants with modified responses to external  
PT signals.  
XX  
PS Claim 1; Page 426-427; 527pp; English.  
XX  
CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide  
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or  
CC pine (Pinus radiata also known as Monterey pine). The protein sequences  
CC are involved in cell signalling. The polynucleotide and protein sequences  
CC can be used to modify the response of plant cells to external signals  
CC e.g. environmental changes or pathogens during the growth and development  
CC  
CC of a plant. They can be used to modify cell proliferation,  
CC differentiation, elongation and survival, resistance to disease and  
CC nutrient metabolism. Examples of modifications which can be produced are  
CC altered fruit ripening and senescence of leaves and flowers e.g. to delay  
CC senescence and prolong the life of cut flowers or enhance senescence of  
CC reproductive organs to engineer sterile plants. Other modifications can  
CC be used to delay senescence in selected cell types or organs providing  
CC fruit and vegetables which have a longer shelf life between harvest and  
CC consumption, or to decrease branching frequency in forest tree species  
CC giving long stretches of valuable knot-free clear wood which can be used  
CC in solid timber furniture and veneers  
XX  
SQ Sequence 2447 BP; 702 A; 528 C; 603 G; 614 T; 0 U; 0 Other;  
Query Match 93.3%; Score 14; DB 3; Length 2447;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 ACACCTATGTGCAC 15  
DB 156 ACACCTATGTGCAC 143  
RESULT 17  
ADQ84612/C  
ID ADQ84612 standard; cDNA; 2757 BP.  
XX  
AC ADQ84612;  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1426.  
XX  
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
KW cancer; cell proliferative disorder; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2004060270-A2.  
XX  
PD 22-JUL-2004.  
XX  
PF 15-OCT-2003; 2003WO-US029126.  
XX  
PR 18-OCT-2002; 2002US-0418988P.  
XX  
PA (GETH ) GENENTECH INC.  
PA (WUTD/) WU T D.  
PA (ZHOU/) ZHOU Y.  
XX  
PI Wu TD, Zhou Y;  
XX  
XX WPI; 2004-534300/51.  
XX  
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,  
PT preventing or treating cell proliferative disorders such as cancer.  
XX  
PS Claim 1; SEQ ID NO 1426; 5504pp; English.  
XX  
XX The present invention describes an isolated tumour-associated antigenic  
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-  
CC (c). Also described: (1) an expression vector comprising the above  
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)  
CC a process for producing a polypeptide; (4) an isolated polypeptide  
CC comprising: (a) an amino acid sequence encoded by any of the above  
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-  
CC length coding region of the above nucleotide sequences; or (c) a sequence  
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)  
CC an isolated antibody that binds to the above polypeptide; (7) a process

CC for producing the antibody; (8) an isolated oligopeptide that binds to  
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
CC binding organic molecule that binds to the above polypeptide; (10) a  
CC composition of matter comprising the above (chimeric) polypeptide,  
CC antibody, oligopeptide or TAT binding organic molecule, in combination  
CC with a carrier; (11) an article of manufacture comprising a container and  
CC the composition of matter contained within the container; (12) methods of  
CC inhibiting the growth of a cell that expresses the above protein, where  
CC the growth of the cell is at least in part dependent upon a growth  
CC potentiating effect of the above protein; (13) a method of  
CC therapeutically treating a mammal having a cancerous tumour comprising  
CC cells that express the above protein; (14) a method of determining the  
CC presence of a protein in a sample suspected of containing the protein  
CC described above; (15) methods of diagnosing the presence of a tumour in a  
CC mammal; (16) a method for treating or preventing a cell proliferative  
CC disorder associated with increased expression or activity of the above  
CC protein; and (17) a method of binding an antibody, oligopeptide or  
CC organic molecule to a cell that expresses the protein described above.  
CC The TAT sequences have cytostatic activities, and can be used in gene  
CC therapy. The composition and methods are useful for diagnosing,  
CC preventing or treating cancer. The composition is also used for preparing  
CC a medicament for the therapeutic treatment or diagnostic detection of a  
CC cell proliferative disorder or cancer. The present sequence represents a  
CC human TAT cDNA sequence from the present invention.

XX  
SQ Sequence 2757 BP; 689 A; 662 C; 671 G; 735 T; 0 U; 0 Other;  
Query Match 93.3%; Score 14; DB 13; Length 2757;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14  
|||||  
DB 1091 GACACCTATGTGCA 1078

RESULT 18  
ADA53027/c  
ID ADA53027 standard; cDNA; 2758 BP.

XX ADA53027;  
XX  
XX 20-NOV-2003 (first entry)  
XX  
XX Human coding sequence, SEQ ID 595.  
DE  
DE Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
KW Inflammatory disease; osteoporosis; neurological disease; gene; ss.  
XX  
XX Homo sapiens.

XX  
XX EP1293569-A2.  
XX  
XX 19-MAR-2003.  
XX  
XX 21-MAR-2002; 2002EP-00006586.  
XX  
XX 14-SEP-2001; 2001JP-00328381.  
XX  
XX 24-JAN-2002; 2002US-0350435P.  
XX  
XX (HELI-) HELIX RES INST.  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX  
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX  
XX WPI; 2003-395539/38.  
DR P-PSDB; ADA54666.

XX New polynucleotides encoding full-length polypeptides, e.g. secretory  
PT and/or membrane proteins, useful for developing medicines for diseases. in

PT which the gene is involved, or as target molecules for gene therapy.  
XX  
XX Claim 1; SEQ ID NO 595; 205pp; English.

XX  
CC The present invention relates to novel human secretory or membrane  
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
CC ADA54071). The coding sequences are useful in the gene therapy of  
CC diseases caused by abnormalities of the proteins, e.g. cancer,  
CC inflammatory diseases, osteoporosis or neurological disease.  
XX  
SQ Sequence 2758 BP; 691 A; 665 C; 666 G; 736 T; 0 U; 0 Other;

Query Match 93.3%; Score 14; DB 10; Length 2758;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14  
|||||  
DB 1092 GACACCTATGTGCA 1079

RESULT 19  
ADM02813  
ID ADM02813 standard; cDNA; 2843 BP.

XX ADM02813;  
XX AC  
XX 20-MAY-2004 (first entry)  
XX  
XX Human cDNA of the invention SEQ ID NO:1498.

XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.  
XX  
XX Homo sapiens.

XX EP1347046-A1.

XX 24-SEP-2003.

XX 12-APR-2002; 2002EP-00008400.

XX 22-MAR-2002; 2002JP-00137785.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-723558/69.  
DR P-PSDB; ADM05256.

XX New polynucleotides and polypeptides are useful in gene therapy, for  
PT developing a diagnostic marker or medicines for regulating their  
PT expression and activity, or as a target of gene therapy.

XX Claim 1; SEQ ID NO 1498; 305pp; English.

XX The invention relates to a novel human polynucleotide and the encoded  
CC polypeptide. A polynucleotide of the invention may have a use in gene  
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
CC as a primer for synthesizing the polynucleotide or as a probe for  
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
CC useful in gene therapy, for developing a diagnostic marker or medicines  
CC for regulating their expression and activity, or as a target of gene  
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
CC are useful as pharmaceutical agents. The present sequence represents a  
CC cDNA sequence of the invention.

XX Sequence 2843 BP; 530 A; 969 C; 893 G; 451 T; 0 U; 0 Other;

Query Match 93.3%; Score 14; DB 11; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;



Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15  
Db 1116 ACACCTATGTGCAC 1129

RESULT 20  
AAK52302/c  
ID AAK52302 standard; cDNA; 4374 BP.  
XX AC AAK52302;  
XX AC AAK52302;  
XX 06-NOV-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 847.  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation; ss.  
XX Homo sapiens.  
XX WO200157190-A2.  
XX 09-AUG-2001.  
XX 05-FEB-2001; 2001WO-US004098.  
XX 03-FEB-2000; 2000US-00496914.  
XX 27-APR-2000; 2000US-00560875.  
XX 20-JUN-2000; 2000US-00598075.  
XX 19-JUL-2000; 2000US-00620325.  
XX 01-SEP-2000; 2000US-00654936.  
XX 15-SEP-2000; 2000US-00663561.  
XX 20-OCT-2000; 2000US-00693325.  
XX 30-NOV-2000; 2000US-00728422.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX WPI; 2001-476283/51.  
XX P-PSDB; AAM79169.  
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
PT in diagnosis and gene therapy.  
XX Claim 1; Page 2836-2842; 6221pp; English.  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication  
XX SQ Sequence 4374 BP; 809 A; 1425 C; 1150 G; 990 T; 0 U; 0 Other;

Query Match 93.3%; Score 14; DB 4; Length 4374;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15  
Db 4327 ACACCTATGTGCAC 4314

RESULT 21  
AAK52301/c  
ID AAK52301 standard; cDNA; 4455 BP.  
XX AC AAK52301;  
XX 06-NOV-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 846.  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation; ss.  
XX Homo sapiens.  
XX WO200157190-A2.  
XX 09-AUG-2001.  
XX 05-FEB-2001; 2001WO-US004098.  
XX 03-FEB-2000; 2000US-00496914.  
XX 27-APR-2000; 2000US-00560875.  
XX 20-JUN-2000; 2000US-00598075.  
XX 19-JUL-2000; 2000US-00620325.  
XX 01-SEP-2000; 2000US-00654936.  
XX 15-SEP-2000; 2000US-00663561.  
XX 20-OCT-2000; 2000US-00693325.  
XX 30-NOV-2000; 2000US-00728422.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX WPI; 2001-476283/51.  
XX P-PSDB; AAM79168.  
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
PT in diagnosis and gene therapy.  
XX Claim 1; Page 2830-2836; 6221pp; English.  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication  
XX SQ Sequence 4455 BP; 827 A; 1452 C; 1167 G; 1009 T; 0 U; 0 Other;

Query Match 93.3%; Score 14; DB 4; Length 4455;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4408 ACACCTATGTGCAC 4395

RESULT 22

AAC77082/c

ID AAC77082 standard; cDNA; 4580 BP.

XX

AC AAC77082;

DT 08-FEB-2001 (first entry)

XX

DE Human ORFX ORF2637 polynucleotide sequence SEQ ID NO:5273.

XX

Human; open reading frame; ORFX; detection; cytostatic; hepatotropic; vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant; immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antinflammatory; antiviral; antibacterial; antifungal; antirheumatic; antithyroid; antianaemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive; ss.

XX

OS Homo sapiens.

XX

XX WO200058473-A2.

XX

XX 05-OCT-2000.

XX

XX 31-MAR-2000; 2000WO-US008621.

XX

XX 31-MAR-1999; 99US-0127607P.

XX

XX 02-APR-1999; 99US-0127636P.

XX

XX 05-APR-1999; 99US-0127728P.

XX

XX 30-MAR-2000; 2000US-00540763.

XX

XX (CURA-) CURAGEN CORP.

XX

XX Shimkets RA, Leach M;

XX

XX WPI; 2000-602362/57.

XX

XX P-PSDB; AAB42873.

XX

XX Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease.

XX

XX Claim 5; Page 4445-4448; 5507pp; English.

XX

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic; anticonvulsant; antiarthritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antinflammatory; antibacterial; antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to

CC enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX

SQ Sequence 4580 BP; 1156 A; 1128 C; 1141 G; 1155 T; 0 U; 0 Other;

Query Match 93.3%; Score 14; DB 3; Length 4580;

Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14

|||||

Db 2402 GACACCTATGTGCA 2389

RESULT 23

ADJ27336/c

ID ADJ27336 standard; cDNA; 5092 BP.

XX

AC ADJ27336;

DT 20-MAY-2004 (first entry)

XX

DE Human NPC1L1 full length coding sequence.

XX

ss; gene; Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network; plasma membrane; transport signal; promoter; sterol regulated element binding protein 1; SREBP1; binding consensus sequence; transmembrane domain; sterol-sensing domain; SSD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal; cholesterol absorption; serum cholesterol; hyperlipidaemia; atherosclerosis; coronary heart disease; stroke; arteriosclerosis.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT CDS 57..4136

FT /\*tag= a

FT /product= "Human NPC1L1"

XX

XX WO2004009772-A2.

XX

XX 29-JAN-2004.

XX

XX 17-JUL-2003; 2003WO-US022467.

XX

XX 19-JUL-2002; 2002US-0397442P.

XX

XX (SCHE ) SCHERING CORP.

XX

XX Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;

XX

XX WPI; 2004-132945/13.

XX

XX P-PSDB; ADJ27337.

XX

XX GENBANK; AF192522.

XX

XX New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting or identifying antagonists of NPC1L1 for inhibiting intestinal cholesterol absorption in a subject, or for treating elevated serum cholesterol or stroke.

XX

XX Example 7; SEQ ID NO 43; 125pp; English.

XX

XX This sequence encodes a Niemann-Pick disease, type C1, gene-like 1 (NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains a motif ADJ27331 which acts as a trans-golgi network to plasma membrane transport signal, and which exhibits limited tissue distribution and gastrointestinal abundance. The human NPC1L1 promoter sequence contains a sterol regulated element binding protein 1 (SREBP1) binding consensus sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-sensing domain (SSD) which is involved in sensing cholesterol levels, possibly by a mechanism which involves direct cholesterol binding. NPC1L1 has 42% amino acid sequence homology to human NPC1, a receptor responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or polynucleotides are useful for detecting or identifying antagonists of

CC NPC1L1, which can be used for inhibiting intestinal cholesterol  
CC absorption in a subject, or for treating medical conditions including  
CC elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary  
CC heart disease, stroke or arteriosclerosis.

XX SQ Sequence 5092 BP; 984 A; 1634 C; 1328 G; 1146 T; 0 U; 0 Other;

Query Match 93.3%; Score 14; DB 12; Length 5092;

Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15

|||||

Db 5024 ACACCTATGTGCAC 5011

RESULT 24

AAK82269

ID AAK82269 standard; DNA; 5180 BP.

XX AAK82269;

XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37081.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 14-AUG-2000; 2000US-0225759P.

XX 18-AUG-2000; 2000US-0226279P.

XX 22-AUG-2000; 2000US-0226681P.

XX 22-AUG-2000; 2000US-0226686P.

XX 23-AUG-2000; 2000US-0227182P.

XX 23-AUG-2000; 2000US-0227009P.

XX 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.

```
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249220P.
PR 17-NOV-2000; 2000US-0249221P.
PR 17-NOV-2000; 2000US-0249222P.
PR 17-NOV-2000; 2000US-0249223P.
PR 17-NOV-2000; 2000US-0249224P.
PR 17-NOV-2000; 2000US-0249225P.
PR 17-NOV-2000; 2000US-0249226P.
PR 17-NOV-2000; 2000US-0249227P.
PR 17-NOV-2000; 2000US-0249228P.
PR 17-NOV-2000; 2000US-0249229P.
PR 17-NOV-2000; 2000US-0249230P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254037P.
PR 05-JAN-2001; 2001US-02559678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides.
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 37081; 3071bp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 5180 BP; 882 A; 1763 C; 1722 G; 813 T; 0 U; 0 Other;
XX
XX Query Match 93.3%; Score 14; DB 4; Length 5180;
XX Best Local Similarity 100.0%; Pred. No. 4.9e+02;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 2 ACACCTATGTGCAC 15
XX |||||||
XX 1875 ACACCTATGTGCAC 1888
XX
XX RESULT 25
XX ACF79823/c
XX ID ACF79823 standard; DNA; 21501 BP.
XX
XX ACF79823;
XX
```

---

```
XX 15-JAN-2004 (first entry)
XX Human SOST gene encoding sclerostin.
XX SOST; sclerostin; osteopathic; cytostatic; gene therapy; human; gene; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Promoter 7960..11960
XX /*tag= a
XX CDS 11961..15360
XX /*tag= b
XX /*product= "Human sclerostin"
XX /*note= "contains an intron"
XX intron 12180..14938
XX /*tag= c
XX WO2003073991-A2.
XX
XX 12-SEP-2003.
XX
XX 28-FEB-2003; 2003WO-US006193.
XX
XX 01-MAR-2002; 2002US-0361258P.
XX 27-AUG-2002; 2002US-0406171P.
XX 13-FEB-2003; 2003US-0447393P.
XX
XX (CELL-) CELTECH R & D INC.
XX (SUTH/) KUNG SUTHERLAND M S.
XX (GEOG/) GEOGHEGAN J C.
XX (YUCC/) YU C.
XX (LATH/) LATHAM J.
XX
XX Kung Sutherland MS, Geoghegan JC, Yu C, Latham J;
XX WPI; 2003-731645/69.
XX P-PSDB; ABR84505.
XX
XX Composition useful for modulating SOST gene expression in mammal and
XX increasing bone density comprises either steroid, prostaglandin, bile
XX salt or nucleotides.
XX
XX Claim 47; Page 121-130; 74pp; English.
XX
XX The present sequence is that of the human SOST gene encoding sclerostin,
XX a bone morphogenetic protein antagonist and regulator of bone matrix
XX formation, which plays a role in the apoptosis of bone progenitor cells.
XX The invention provides antagonists to the sclerostin protein and methods
XX for identifying new sclerostin antagonists. It also provides molecules
XX that can depress expression of the gene, such as small interfering RNA
XX molecules (siRNA, see ACF79799-ACF79810). Pharmaceutical compositions
XX comprising such molecules are useful for increasing bone density in a
XX mammal and for decreasing apoptosis of bone cells (claimed). They can
XX also be used to modulate bone resorption and augment bone mineralization,
XX to prevent or treat osteoporosis of bone-related cells or loss of bone
XX density, and to treat osteoporosis, osteomyelitis, hypercalcaemia,
XX osteopenia brought on by surgery or steroid administration. Paget's
XX disease, osteonecrosis, bone loss due to rheumatoid arthritis,
XX periodontal bone loss, prosthetic loosening and osteolytic metastasis
XX
XX Sequence 21501 BP; 5546 A; 5259 C; 5419 G; 5277 T; 0 U; 0 Other;
XX
XX Query Match 93.3%; Score 14; DB 10; Length 21501;
XX Best Local Similarity 100.0%; Pred. No. 5.6e+02;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GACACCTATGTGCA 14
XX |||||||
XX 2535 GACACCTATGTGCA 2522
XX
XX
```

```
RESULT 26
ADQ97220
ID ADQ97220 standard; DNA; 52640 BP.
XX AC
XX ADQ97220;
XX DT
XX 07-OCT-2004 (first entry)
XX DE
XX Mouse cancer associated sequence MD08-015, SEQ ID 196.
XX KW
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX OS
XX Mus musculus.
XX PN
XX WO2004060304-A2.
XX PD
XX 22-JUL-2004.
XX PF
XX 22-DEC-2003; 2003WO-US041389.
XX PR
XX 27-DEC-2002; 2002US-00330773.
XX PA
XX (SAGR-) SAGRES DISCOVERY INC.
XX PI
XX Morris DW, Malandro MS;
XX XX
XX WIPI; 2004-543781/52.
XX XX
XX New isolated cancer associated nucleic acids comprising at least 10
XX PT
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX PT
XX cancers such as leukemia and lymphoma.
XX XX
XX Claim 1; SEQ ID NO 196; 199pp; English.
XX XX
XX The present invention relates to cancer associated sequences (ADQ97023-
XX CC
XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX CC
XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX CC
XX data for this patent did not form part of the printed specification, but
XX CC
XX was obtained in electronic format directly from WIPO at
XX CC
XX ftp.wipo.int/pub/published_pct_sequences.
XX XX
XX Sequence 52640 BP; 14760 A; 11829 C; 12664 G; 12714 T; 0 U; 673 Other;
XX SQ
Query Match 93.3%; Score 14; DB 12; Length 52640;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACACCTATGTGCA 14
DB 23120 GACACCTATGTGCA 23133
RESULT 27
AAF22295/c
ID AAF22295 standard; DNA; 80450 BP.
XX AC
XX AAF22295;
XX DT
XX 20-MAR-2001 (first entry)
XX DE
XX BAC containing repeats from centromeres 1-4 #18.
XX KW
XX Centromere; microsome; vector; ds.
XX OS
XX Arabidopsis thaliana.
XX XX
XX WO200055325-A2.
XX PN
XX 21-SEP-2000.
XX PD
XX 17-MAR-2000; 2000WO-US007392.
XX PF
XX 18-MAR-1999; 99US-0125219P.
XX PR
01-APR-1999; 99US-0127409P.
PR 13-MAY-1999; 99US-0134770P.
PR 13-SEP-1999; 99US-0153584P.
PR 17-SEP-1999; 99US-0154603P.
PR 16-DEC-1999; 99US-0172493P.
XX XX
XX (UYCH-) UNIV CHICAGO.
XX PA
XX Preuss D, Copenhaver G, Keith K;
XX PI
XX WIPI; 2000-587529/55.
XX DR
XX Recombinant DNA construct comprising a plant centromere, useful for
XX PT
XX producing stably inherited microsome which can serve as vectors for the
XX PT
XX construction of transgenic plant and animal cells expressing selected
XX XX
XX Claim 102; Page 668-686; 1449pp; English.
XX PS
XX The present invention relates to a recombinant DNA construct of a plant
XX CC
XX (Arabidopsis thaliana) centromere. The constructs are useful for
XX CC
XX producing stably inherited microsome which can serve as vectors for the
XX CC
XX construction of transgenic plant and animal cells expressing selected
XX CC
XX proteins such as hormones, enzymes, interleukins, clotting factors,
XX CC
XX cytokines, antibodies, and growth factors
XX XX
XX Sequence 80450 BP; 25014 A; 14662 C; 14629 G; 26145 T; 0 U; 0 Other;
XX SQ
Query Match 93.3%; Score 14; DB 3; Length 80450;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACACCTATGTGCA 14
DB 1764 GACACCTATGTGCA 1751
RESULT 28
ADF11646
ID ADF11646 standard; DNA; 94752 BP.
XX AC
XX ADF11646;
XX DT
XX 12-FEB-2004 (first entry)
XX DE
XX Human chromosome 17 clone HPRC905N1 nucleic acid.
XX XX
XX osteopathic; Gene therapy; bone mineral density; sclerostin gene region;
XX KW
XX osteoporosis; osteopenia; bone dysplasia; bone fracture; chromosome 17;
XX KW
XX ss.
XX XX
XX Homo sapiens.
XX OS
XX WO2003087763-A2.
XX PN
XX 23-OCT-2003.
XX PD
XX 03-APR-2003; 2003WO-US010649.
XX PF
XX 03-APR-2002; 2002US-0370088P.
XX PR
XX (CELL-) CELLSCH R & D INC.
XX PA
XX (UYRO-) UNIV ROTTERDAM ERASMUS.
XX XX
XX Brunkow ME, Charmley PR, Proll S, Paepers BW, Uitterlinden AG;
XX PI
XX WIPI; 2003-833790/77.
XX DR
XX Determining a risk for or presence of altered bone mineral density (e.g.
XX PT
XX osteoporosis) in a subject comprises determining the presence or absence
XX PT
XX of a sclerostin gene region nucleotide polymorphism in a biological
XX PT
XX sample from a subject.
XX XX
XX Disclosure; SEQ ID NO 34; 114pp; English.
XX PS
```

XX The invention relates to a method of determining a risk for or presence  
 CC of altered bone mineral density (BMD) in a subject by determining the  
 CC presence or absence of at least one sclerostin gene region nucleotide  
 CC polymorphism in a biological sample from a subject where the presence of  
 CC at least one polymorphism at a position that corresponds to a non-coding  
 CC region of the 130320 bp sclerostin gene region (SOST) indicates an  
 CC increased risk of altered BMD. The composition and methods are useful in  
 CC determining a subject a risk for having, or presence of, altered bone  
 CC mineral density, such as osteoporosis, osteopenia, bone dysplasia, bone  
 CC fracture or other conditions characterized by decreased or increased bone  
 CC density. These may also be used in identifying agents that may be used  
 CC for treating the above diseases, disorders or conditions associated with  
 CC altered BMD. In addition, these may be used for pharmacogenomic purposes,  
 CC e.g. to stratify patient populations according to suitability of a  
 CC particular therapeutic agent for use in the population. This sequence  
 CC corresponds to the human chromosome 17 clone HPRC905N1.  
 XX  
 SQ Sequence 94752 BP; 24324 A; 23247 C; 22257 G; 24924 T; 0 U; 0 Other;  
 Query Match 93.3%; Score 14; DB 10; Length 94752;  
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACACCTATGTGCA 14  
 Db 18695 GACACCTATGTGCA 18708  
 RESULT 29  
 AAD53223\_0/c  
 WP Sequence split into 5 fragments LOCUS AAD53223 Accession Aad53223  
 WP Fragment Name Begin End  
 WP AAD53223\_0 1 110000  
 WP AAD53223\_1 100001 210000  
 WP AAD53223\_2 200001 310000  
 WP AAD53223\_3 300001 410000  
 WP AAD53223\_4 400001 487980  
 ID AAD53223 standard; DNA; 487980 BP.  
 XX  
 AC AAD53223;  
 XX  
 DT 28-MAY-2003 (first entry)  
 DE Human chromosome 3 p-arm breakpoint region.  
 XX  
 KW Behavioural disorder; attention deficit hyperactivity disorder; ADHD;  
 KW molecular marker; intellectual disorder; gene therapy; chromosome 3p;  
 KW human; db.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200290541-A1.  
 XX  
 PD 14-NOV-2002.  
 XX  
 PF 03-MAY-2002; 2002WO-AU000556.  
 XX  
 PR 03-MAY-2001; 2001AU-00004756.  
 PR 04-JUN-2001; 2001AU-00005426.  
 PR 04-JUN-2001; 2001US-0295811P.  
 XX  
 PA (MURD-) MURDOCH CHILDRENS RES INST.  
 PA (DELA/) DELATYCKI M.  
 XX  
 PI Williamson R, Dahl HM, Forrest SM, Wilcox SA, De Silva MG;  
 PI Elliott KS, Lynch M;  
 XX  
 DR WPI; 2003-111974/10.  
 XX  
 PT New molecular marker of a behavioral disorder, useful for diagnosing  
 PT behavioral disorder, or assessing the likelihood of developing behavioral  
 PT disorder, e.g. Attention Deficit Hyperactivity Disorder or intellectual

PT disorders.  
 XX Claim 20; Fig 8; 390pp; English.  
 XX  
 CC The invention relates to a molecular marker of a behavioural disorder,  
 CC which is in a genetic form, and comprises a genetic location on  
 CC chromosome 3 or an equivalent location on another chromosome, where a  
 CC mutation at the location alone or in combination with environmental or  
 CC other genetic factors is associated with or otherwise facilitates the  
 CC development or progression of the behavioural disorder. The molecular  
 CC marker is useful for diagnosing behavioural disorder, or assessing the  
 CC likelihood of developing behavioural disorder, e.g. attention deficit  
 CC hyperactivity disorder (ADHD) or intellectual disorders. They are also  
 CC useful for facilitating the development of therapeutic protocols for  
 CC treatment of the behavioural disorders. Sequences of the invention are  
 CC useful in manufacturing a genetic probe to determine the likelihood of a  
 CC subject having a behavioural disorder, such as ADHD. They are also useful  
 CC for diagnosing, preventing or treating a behavioural disorder. The  
 CC invention is useful in gene therapy. The present sequence is human  
 CC chromosome 3 p-arm breakpoint region from clones BAC RP11-804h8, 646d/13,  
 CC 73i17, 3f4 and 15id23. This sequence is used in the invention  
 XX  
 SQ Sequence 487980 BP; 127777A; 93388C; 96796G; 140865T; 0U; 29154Other;  
 Query Match 93.3%; Score 14; DB 8; Length 110000;  
 Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 ACACCTATGTGCAC 15  
 Db 11820 ACACCTATGTGCAC 11807  
 RESULT 30  
 ADF11613/C  
 ID ADF11613 standard; DNA; 130320 BP.  
 XX  
 AC ADF11613;  
 XX  
 DT 12-FEB-2004 (first entry)  
 DE Human sclerostin gene region.  
 XX  
 KW ds; osteopathic; gene therapy; bone mineral density;  
 KW sclerostin gene region; osteoporosis; osteopenia; bone dysplasia;  
 KW bone fracture.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT variation replace(4103,G)  
 FT /\*tag= a  
 FT variation replace(10357,T)  
 FT /\*tag= b  
 FT variation replace(10565, .10566,AGGAC)  
 FT /\*tag= c  
 FT variation replace(17966,G)  
 FT /\*tag= d  
 FT variation replace(18293,G)  
 FT /\*tag= e  
 FT variation replace(58083,C)  
 FT /\*tag= f  
 FT variation replace(74235,G)  
 FT /\*tag= g  
 FT variation replace(91068,G)  
 FT /\*tag= h  
 XX  
 PN W02003087763-A2.  
 XX  
 PD 23-OCT-2003.  
 XX  
 PF 03-APR-2003; 2003WO-US010649.  
 XX

PR 03-APR-2002; 2002US-0370088P.  
XX (CELL-) CELLTech R & D INC.  
PA (UTRO-) UNIV ROTTERDAM ERASMUS.  
XX Brunkow ME, Charmley PR, Proll S, Paepers BW, Uitterlinden AG;  
XX WPI; 2003-833790/77.  
DR  
XX  
XX Determining a risk for or presence of altered bone mineral density (e.g.  
PT osteoporosis) in a subject comprises determining the presence or absence  
PT of a sclerostin gene region nucleotide polymorphism in a biological  
PT sample from a subject.  
XX  
XX Claim 21; SEQ ID NO 1; 114pp; English.  
PS  
XX The invention relates to a method of determining a risk for or presence  
XX of altered bone mineral density (BMD) in a subject by determining the  
CC presence or absence of at least one sclerostin gene region nucleotide  
CC polymorphism in a biological sample from a subject where the presence of  
CC at least one polymorphism at a position that corresponds to a non-coding  
CC region of the 130320 bp sclerostin gene region (SOST) indicates an  
CC increased risk of altered BMD. The composition and methods are useful in  
CC determining a subject a risk for having, or presence of, altered bone  
CC mineral density, such as osteoporosis, osteopenia, bone dysplasia, bone  
CC fracture or other conditions characterized by decreased or increased bone  
CC density. These may also be used in identifying agents that may be used  
CC for treating the above diseases, disorders or conditions associated with  
CC altered BMD. In addition, these may be used for pharmacogenomic purposes,  
CC e.g. to stratify patient populations according to suitability of a  
CC particular therapeutic agent for use in the population. This sequence  
CC corresponds to the human sclerostin gene region.  
XX  
SQ Sequence 130320 BP; 33204 A; 32954 C; 31896 G; 32253 T; 0 U; 13 Other;  
  
Query Match 93.3%; Score 14; DB 10; Length 130320;  
Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GACACCTATGTGCA 14  
DB 2536 GACACCTATGTGCA 2523  
  
RESULT 31  
ADL13861/c  
ID ADL13861 standard; DNA; 154681 BP.  
XX  
XX  
AC ADL13861;  
XX  
XX 06-MAY-2004 (first entry)  
XX  
XX Osteoarthritis-associated polymorphic nucleotide #393.  
XX  
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;  
XX joint space narrowing; osteophyte development; joint pain;  
XX osteoarthritis; SNP; single nucleotide polymorphism.  
XX  
XX Homo sapiens.  
XX  
XX WO2003054166-A2.  
PN  
XX  
XX 03-JUL-2003.  
PD  
XX  
XX 19-DEC-2002; 2002WO-US041225.  
PF  
XX  
XX 20-DEC-2001; 2001US-0342603P.  
PR  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
PA  
XX  
XX Jones KA, Schafer A;  
PI  
XX WPI; 2003-559141/52.  
DR  
  
XX Determining susceptibility of an individual to joint space narrowing,  
PT osteophyte development and/or joint pain comprises identifying whether  
PT the individual has at least one polymorphism in a polynucleotide encoding  
PT a protein.  
XX  
XX Disclosure; SEQ ID NO 393; 297pp; English.  
PS  
XX The invention relates to a method of determining susceptibility of an  
XX individual to joint space narrowing and/or osteophyte development and/or  
CC joint pain comprising identifying whether the individual has at least one  
CC polymorphism in a polynucleotide encoding at least one of the protein  
CC listed in the specification. The methods, composition and agent are  
CC useful for modulating the susceptibility of an individual to joint space  
CC narrowing and/or osteophyte development and/or joint pain that is  
CC associated with a disease, preferably osteoarthritis. The cell line and  
CC the non-human animal are useful for screening for an agent for diagnosing  
CC an individual having susceptibility to joint space narrowing and/or  
CC osteophyte development and/or joint pain. This sequence corresponds to  
CC the polynucleotide encoding a protein listed in the specification. (Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences).  
XX  
SQ Sequence 154681 BP; 41177 A; 36356 C; 36934 G; 40214 T; 0 U; 0 Other;  
  
Query Match 93.3%; Score 14; DB 10; Length 154681;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 ACACCTATGTGCAC 15  
DB 953 ACACCTATGTGCAC 940  
  
RESULT 32  
AAH88704  
ID AAH88704 standard; DNA; 160755 BP.  
XX  
XX  
AC AAH88704;  
XX  
XX 26-FEB-2002 (first entry)  
XX  
XX Human DNA sequence SEQ ID 544.  
DE  
XX  
XX Single nucleotide polymorphism; SNP; biallelic marker; human;  
XX central nervous system disorder; CNS; ds.  
KW  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200151659-A2.  
PN  
XX  
XX 19-JUL-2001.  
PD  
XX  
XX 11-JAN-2001; 2001WO-IB000116.  
PF  
XX  
XX 13-JAN-2000; 2000US-0175854P.  
PR  
XX  
XX (GEST ) GENSET.  
PA  
XX  
XX Chu T, Blumenfeld M, Cohen D;  
PI  
XX  
XX WPI; 2001-483085/52.  
DR  
XX  
XX Isolated polynucleotides, useful for genotyping nucleic acids for  
PT biallelic markers for the diagnosis of depression, comprises central  
PT nervous system disorder related biallelic marker.  
XX  
XX Disclosure; Page 476-519; 519pp; English.  
PS  
XX  
XX The present invention relates to biallelic markers derived from human  
CC genes involved in central nervous system (CNS) disorders (see AAH88161-  
CC AAH88702). The markers have a single nucleotide polymorphism (SNP) and

CC are useful in determining the genetic predisposition of individuals to  
CC CNS disorders, by identifying the nucleotides at a set of genetic markers  
CC in a biological sample, where the markers comprise at least one CNS  
CC disorder related marker. The present sequence was used to illustrate the  
CC invention  
XX  
SQ Sequence 160755 BP; 43057 A; 36601 C; 38189 G; 42727 T; 0 U; 181 Other;  
Query Match 93.3%; Score 14; DB 4; Length 160755;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACACCTATGTGCA 14  
|||||  
Db 30295 GACACCTATGTGCA 30308  
RESULT 33  
ADQ97348/c  
ID ADQ97348 standard; DNA; 204803 BP.  
XX AC ADQ97348;  
XX 07-OCT-2004 (first entry)  
XX Mouse cancer associated sequence MP08-036, SEQ ID 325.  
DE  
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.  
XX Mus musculus.  
XX  
PN W02004060304-A2.  
XX 22-JUL-2004.  
XX 22-DEC-2003; 2003WO-US041389.  
XX 27-DEC-2002; 2002US-00330773.  
XX (SAGR-) SAGRES DISCOVERY INC.  
PA Morris DW, Malandro MS;  
PI WPI; 2004-543781/52.  
XX  
XX New isolated cancer associated nucleic acids comprising at least 10  
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating  
PT cancers such as leukemia and lymphoma.  
XX  
PS Claim 1; SEQ ID NO 325; 199pp; English.  
XX  
CC The present invention relates to cancer associated sequences (ADQ97025-  
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or  
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 204803 BP; 55167 A; 44951 C; 46757 G; 55510 T; 0 U; 2418 Other;  
Query Match 93.3%; Score 14; DB 12; Length 204803;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 ACACCTATGTGCAC 15  
|||||  
Db 76954 ACACCTATGTGCAC 76941  
RESULT 34  
ACN44418  
ID ACN44418 standard; DNA; 213040 BP.  
XX

AC ACN44418;  
XX 18-NOV-2004 (first entry)  
DT  
XX Human genomic sequence hCG41574.  
DE  
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.  
XX Homo sapiens.  
OS  
XX W02003073826-A2.  
PN  
XX 12-SEP-2003.  
PD  
XX 28-FEB-2003; 2003WO-US006235.  
PF  
XX 01-MAR-2002; 2002US-00087192.  
PR  
XX (SAGR-) SAGRES DISCOVERY.  
PA Morris DW;  
XX  
PI WPI; 2003-328604/31.  
DR  
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
PT comprises a nucleotide sequence.  
PT  
XX Claim 1; SEQ ID NO 856; Opp; English.  
PS  
XX The present invention relates to novel DNA and protein sequences which  
CC are associated with carcinomas. The sequences are useful for: (i) for  
CC screening drug candidates; (ii) for screening of bioactive agent capable  
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
CC a bioactive agent capable of modulating the activity of CAP; (iv) for  
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
CC determining Carcinoma Associated (CA) gene copy number. In addition, the  
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of  
CC carcinoma including lymphoma. The present sequence is one such CA coding  
CC sequence. Note: This patent is an equivalent to basic patent  
CC US2002182586A1, for which no sequence data was published  
XX  
SQ Sequence 213040 BP; 60152 A; 41111 C; 42756 G; 63937 T; 0 U; 5084 Other;  
Query Match 93.3%; Score 14; DB 11; Length 213040;  
Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACACCTATGTGCA 14  
|||||  
Db 61402 GACACCTATGTGCA 61415  
RESULT 35  
ADL13809/c  
ID ADL13809 standard; DNA; 214019 BP.  
XX  
AC ADL13809;  
XX  
XX 06-MAY-2004 (first entry)  
DT  
XX Osteoarthritis-associated polymorphic nucleotide #341.  
DE  
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;  
KW Joint space narrowing; osteophyte development; joint pain;  
KW osteoarthritis; SNP; single nucleotide polymorphism.  
XX  
XX Homo sapiens.  
OS  
XX W02003054166-A2.  
PN  
XX



PD 03-JUL-2003.  
XX 19-DEC-2002; 2002WO-US041225.  
XX 20-DEC-2001; 2001US-0342603P.  
XX (INCYTE GENOMICS INC.  
XX Jones KA, Schafer A;  
XX WPI; 2003-559141/52.  
XX  
XX Determining susceptibility of an individual to joint space narrowing,  
XX osteophyte development and/or joint pain comprises identifying whether  
XX the individual has at least one polymorphism in a polynucleotide encoding  
XX a protein.  
XX Disclosure; SEQ ID NO 341; 297pp; English.  
XX  
XX The invention relates to a method of determining susceptibility of an  
XX individual to joint space narrowing and/or osteophyte development and/or  
XX joint pain comprising identifying whether the individual has at least one  
XX polymorphism in a polynucleotide encoding at least one of the protein  
XX listed in the specification. The methods, composition and agent are  
XX useful for modulating the susceptibility of an individual to joint space  
XX narrowing and/or osteophyte development and/or joint pain that is  
XX associated with a disease, preferably osteoarthritis. The cell line and  
XX the non-human animal are useful for screening for an agent for diagnosing  
XX an individual having susceptibility to joint space narrowing and/or  
XX osteophyte development and/or joint pain. This sequence corresponds to  
XX the polynucleotide encoding a protein listed in the specification. (Note:  
XX The sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences).  
XX  
XX Sequence 214019 BP; 66854 A; 41152 C; 39926 G; 65683 T; 0 U; 404 Other;  
XX  
XX Query Match 93.3%; Score 14; DB 10; Length 214019;  
XX Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 2 ACACCTATGTGCAC 15  
XX 14606 ACACCTATGTGCAC 14593  
XX  
XX  
XX RESULT 36  
XX ABD32694/C  
XX ID ABD32694 standard; DNA; 252907 BP.  
XX AC ABD32694;  
XX  
XX 18-NOV-2004 (first entry)  
XX DT  
XX DE Human cancer-associated genomic DNA HD14-012.  
XX  
XX Human; db; cancer-associated protein; gene; cytostatic; cancer;  
XX KW leukaemia; lymphoma; CAP.  
XX OS Homo sapiens.  
XX PN W02004074320-A2.  
XX PD  
XX 02-SEP-2004.  
XX PF  
XX 17-FEB-2004; 2004WO-US004730.  
XX PR  
XX 14-FEB-2003; 2003US-00367094.  
XX PR 14-MAR-2003; 2003US-00388838.  
XX PR 15-APR-2003; 2003US-00417375.  
XX PR 13-JUN-2003; 2003US-00461862.  
XX PR 15-SEP-2003; 2003US-00663431.  
XX PR 15-DEC-2003; 2003US-00737318.  
XX  
XX (SAGR-) SAGRES DISCOVERY INC.  
XX Morris DW, Morris DW, Malandro MS;  
XX WPI; 2004-652914/63.  
XX  
XX New isolated cancer-associated polynucleotides and polypeptides useful  
XX for diagnosing, preventing or treating cancers, especially lymphoma and  
XX leukemia, or in screening for agents that modulate cancer.  
XX  
XX claim 16; seqid 255; 310pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising at least 10  
XX contiguous nucleotides of any of the 233 polynucleotide sequences given  
XX in the specification, or its complement. The nucleic acids encode cancer-  
XX associated proteins. Also included are an expression vector comprising  
XX the isolated nucleic acid cited above, a host cell comprising the above  
XX recombinant nucleic acid or expression vector, a microarray for detecting  
XX a cancer-associated (CA) nucleic acid comprising at least one probe  
XX comprising at least 10 contiguous nucleotides of any of the above-  
XX mentioned nucleotide sequences, an isolated polypeptide (encoded within  
XX an open reading frame of a CA sequence selected from any of the 95  
XX polynucleotide sequences as mentioned in the specification, or its  
XX complement), an isolated antibody, (or its antigen binding fragment) that  
XX binds to the above polypeptide, a hybridoma that produces the above  
XX monoclonal antibody, a pharmaceutical composition comprising the above  
XX antibody and a pharmaceutical excipient, a kit for detecting cancer  
XX cells (comprising the antibody cited above, methods for diagnosing cancer  
XX or for detecting the presence or absence of cancer cells in an  
XX individual, a method for inhibiting growth of cancer cells in an  
XX individual, a method for delivering a therapeutic agent to cancer cells  
XX in an individual, an electronic library comprising the above  
XX polynucleotide or polypeptide (or their fragments), methods of screening  
XX for anticancer activity or for a bioactive agent capable of modulating  
XX the activity of a CA protein (CAP), methods for detecting cancer  
XX associated with expression of a polypeptide in a test cell sample, a  
XX method for treating cancers and a method for inhibiting the expression of  
XX CA gene in a cell. The composition and methods are useful for detecting,  
XX diagnosing, preventing and treating cancers, especially lymphoma and  
XX leukaemia. These may also be used in screening for agents that modulate  
XX cancer. The present sequence is a human CAP genomic sequence. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Query Match 93.3%; Score 14; DB 13; Length 252907;  
XX Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1 GACACCTATGTGCA 14  
XX 72693 GACACCTATGTGCA 72680  
XX  
XX  
XX RESULT 37  
XX AAQ43233  
XX ID AAQ43233 standard; DNA; 24 BP.  
XX AC AAQ43233;  
XX  
XX 25-MAR-2003 (revised)  
XX DT 13-OCT-1993 (first entry)  
XX DX  
XX DE B-B10 V region primer VH2.  
XX  
XX Complementarity-determining region; CDR; humanised; antibody; hIL2R;  
XX KW human; interleukin; IL-2; receptor; murine; anti-human; Ab; T-cell;  
XX KW monoclonal antibody; B-B10; mixed lymphocyte reaction; variable; V;  
XX KW region; PCR; framework; plasmid; heavy; H; light; L; amplify; primer;  
XX KW polymerase chain reaction; ss.

```

XX OS Synthetic.
XX PN WO9311238-A1.
XX PD 10-JUN-1993.
XX PF 03-DEC-1992; 92WO-JP001583.
XX PR 06-DEC-1991; 91JP-00323319.
XX PA (SUMU) SUMITOMO PHARM CO LTD.
XX PA (BIOT) BIOTEST PHARMA GMBH.
XX PA (INNO-) INNOTHERAPIE LAB.
XX PI Nakatani T, Gomi H, Wijdenes J, Noguchi H;
XX WI; 1993-197057/24.
XX PT Humanised antibody comprising - CDR region of mouse MAB B-B10 specific
XX PT for IL-2 receptor useful for treating carcinoma expressing IL-2 receptor.
XX PS Disclosure; Page 46; 62pp; English.
XX CC The sequences given in AAQ43233-36 are primers which were used in the
XX CC cloning of DNA encoding the variable (V) regions of the murine anti-
XX CC human IL-2 receptor monoclonal Ab (MAB) B-B10. This MAB was used in the
XX CC construction of a humanised antibody (Ab) which binds specifically to
XX CC human interleukin (IL)-2 receptor (hIL2R). The complementarity-
XX CC determining regions (CDRs) for the hIL2R MAB were derived from B-B10 (see
XX CC also AARJ7599-04). The hIL2R MAB is antagonistic to the binding of IL-2
XX CC to the IL-2 receptor on human T-cells. It also inhibits the human mixed
XX CC lymphocyte reaction. The cDNA encoding the variable (V) region of the B-
XX CC B10 Ab was cloned by PCR and sequenced (see also AAQ43233-36) A human Ab
XX CC with high levels of amino acid sequence homology to the murine sequence
XX CC was selected and the framework of this Ab was bound with the B-B10 V
XX CC region CDR and a part of the framework to design several kinds of the
XX CC humanised B-B10 V region. The DNA sequence coding this humanised B-B10
XX CC was synthesised and a plasmid expressing humanised B-B10 was constructed.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 24 BP; 11 A; 6 C; 2 G; 5 T; 0 U; 0 Other;

Query Match 89.3%; Score 13.4; DB 2; Length 24;
Best Local Similarity 93.3%; Pred. No. 6.9e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCAC 15
Db ||||| ||||| |||||
10 GACACCTATATGCAC 24

RESULT 38
ABZ68657
ID ABZ68657 standard; DNA; 45 BP.
XX AC ABZ68657;
XX DT 16-MAY-2003 (first entry)
XX DE Primer used to humanise K121 antibody heavy chain variable region.
XX KW K121 antibody; K121-like antibody; kappa-type myeloma cell;
XX KW kappa-type multiple myeloma; haematopoietic cell transplantation;
XX KW apoptosis; kappa myeloma antigen; primer; ss.
XX OS Synthetic.
XX PN WO2003004056-A1.
XX PD 16-JAN-2003.
XX PF 05-JUL-2002; 2002WO-AU000896.

06-JUL-2001; 2001AU-00006179.
(PACM-) PACWAB PTY LTD.
Raison RL, Dunn RD, Choo BHA;
WPI; 2003-210317/20.
Treating kappa-type multiple myeloma in a subject by administering a K121
-like antibody not conjugated to a toxin or a cytolytic agent.
Example 15; Fig 15b; 65pp; English.
Primers ABZ68657-66 were used to humanise the murine K121 antibody heavy
chain variable region. The specification describes a K121-like antibody
that competes with K121 for binding to kappa-type myeloma cells. The K121
-like antibody is used in the method of the invention. The specification
describes a method for treating kappa-type multiple myeloma in a subject,
comprising administering a K121-like antibody which is not conjugated to
a toxin or a cytolytic agent. The method is useful for treating kappa-
type multiple myeloma, autologous haematopoietic cell transplantation,
killing kappa-type myeloma cells in a mixed population of cells and
inducing apoptosis in kappa myeloma antigen (KMA) bearing cells
Sequence 45 BP; 11 A; 14 C; 11 G; 9 T; 0 U; 0 Other;

Query Match 89.3%; Score 13.4; DB 10; Length 45;
Best Local Similarity 93.3%; Pred. No. 7.3e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCAC 15
Db ||||| ||||| |||||
11 GACACCTATATGCAC 25

RESULT 39
ADQ31301
ID ADQ31301 standard; DNA; 53 BP.
XX AC ADQ31301;
XX DT 09-SEP-2004 (first entry)
XX DE Mutagenic PCR primer to humanise murine 11K2 heavy chain antibody 2 ID67.
XX KW ss; murine; 11K2; monocyte chemotactic protein; MCP;
XX KW beta-chemokine family; MCP-3; glomerulonephritis; scleroderma; cirrhosis;
XX KW multiple sclerosis; lupus nephritis; atherosclerosis;
XX KW inflammatory bowel disease; rheumatoid arthritis; inflammatory disease;
XX KW fibrotic disorder; cancer; immunopathological disorder;
XX KW antiarteriosclerotic; antiarthritic; immunopathological disorder;
XX KW cytostatic; dermatological; hepatotropic; immunomodulatory; antirheumatic;
XX KW neuroprotective; PCR; primer; mouse; mutagenic.
XX OS Mus musculus.
XX OS Synthetic.
XX PN WO2004050836-A2.
XX PD 17-JUN-2004.
XX PF 25-NOV-2003; 2003WO-US037834.
XX PR 27-NOV-2002; 2002US-0430007P.
XX PA (BIOG-) BIOGEN IDEC MA INC.
XX PI De Fougères AR, Koteliński VE, Garber E, Reid C, Saldanha JW;
XX PI Van Vlijmen H;
XX WPI; 2004-461110/43.

```

PT New antibodies against monocyte chemotactic proteins (MCP), useful for  
PT treating or preventing disorders associated with detrimental MCP  
PT activity, e.g. glomerulonephritis, scleroderma, multiple sclerosis, or  
PT atherosclerosis.  
XX Example 14; SEQ ID NO 67; 200pp; English.  
XX  
XX This invention relates to an antibody for treating or preventing  
CC disorders associated with detrimental monocyte chemotactic protein (MCP)  
CC activity. Specifically, it refers to humanised antibodies that bind to  
CC members of the beta-chemokine family (of which MCP-1, MCP-2 and MCP-3  
CC belong) and in particular antibodies that have been modelled on, and  
CC modified from, the variable complementarity determining regions (CDRs) of  
CC the murine 11K2 and 1A1 immunoglobulin sequences. The present invention  
CC describes using these antibodies to treat or prevent diseases and  
CC disorders including glomerulonephritis, scleroderma, cirrhosis, multiple  
CC sclerosis, lupus nephritis, atherosclerosis, inflammatory bowel diseases,  
CC rheumatoid arthritis, inflammatory diseases, fibrotic disorders, cancer  
CC and immunopathological disorders. Accordingly, they can be used in the  
CC development of pharmaceutical compositions that exhibit  
CC antiarteriosclerotic, antiarthritic, antiinflammatory, antirheumatic,  
CC cytostatic, dermatological, hepatotropic, immunomodulator, nephrotropic  
CC and neuroprotective activities. This oligonucleotide sequence is a  
CC mutagenic PCR primer used to generate humanised 11K2 antibodies of the  
CC invention.  
XX  
SQ Sequence 53 BP; 11 A; 17 C; 13 G; 12 T; 0 U; 0 Other;  
Query Match 89.3%; Score 13.4; DB 12; Length 53;  
Best Local Similarity 93.3%; Pred. No. 7.4e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GACACCTATCTGCAC 15  
Db 35 GACACCTATATGCAC 49  
RESULT 40  
ADI25193/c  
ID ADI25193 standard; DNA; 54 BP.  
XX  
XX AC ADI25193;  
XX  
XX DT 22-APR-2004 (first entry)  
XX  
XX DE Humanised murine HP1/2 Ig VH mutagenesis oligonucleotide #1.  
XX  
XX KW humanised recombinant antibody; alpha4-binding fragment; inflammation;  
KW VCAM-1; VLA-4; adhesion pathway; asthma; inflammatory bowel disease;  
KW murine; HP1/2; Ig; immunoglobulin; VH; mouse; human; mutagenesis; ss.  
XX  
XX OS Mus sp.  
OS Homo sapiens.  
OS Chimeric.  
XX  
XX PN US6602503-B1.  
XX  
XX PD 05-AUG-2003.  
XX  
XX PF 31-MAY-1995; 95US-00454899.  
XX  
XX PR 12-JAN-1993; 93US-00004798.  
XX PR 07-JAN-1994; 94WO-US000266.  
XX  
XX PA (BIOJ ) BIOGEN INC.  
XX  
XX PI Lobb RR, Carr FJ, Tempest PR;  
XX WPI; 2003-895269/82.  
XX  
XX PT New humanized recombinant antibody molecule or its alpha4-binding  
PT fragment, useful for diagnosing or treating inflammation associated with  
PT a VCAM-1/VLA-4 adhesion pathway in a mammal e.g., asthma or inflammatory

PT bowel disease.  
XX Example 3; SEQ ID NO 18; 60pp; English.  
XX  
XX The invention relates to a new humanised recombinant antibody molecule or  
CC its alpha4-binding fragment comprises at least one antibody heavy or  
CC light chain. The humanised recombinant antibody molecule is useful for  
CC diagnosing or treating inflammation associated with a VCAM-1/VLA-4  
CC adhesion pathway in a mammalian subject, e.g. asthma or inflammatory  
CC bowel disease. The present sequence represents a humanised murine HP1/2  
CC Ig VH mutagenesis oligonucleotide.  
XX  
SQ Sequence 54 BP; 13 A; 13 C; 13 G; 15 T; 0 U; 0 Other;  
Query Match 89.3%; Score 13.4; DB 10; Length 54;  
Best Local Similarity 93.3%; Pred. No. 7.4e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GACACCTATCTGCAC 15  
Db 26 GACACCTATATGCAC 12  
Search completed: June 4, 2005, 13:53:54  
Job time : 94.4242 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 13:25:28 ; Search time 29.5455 Seconds  
(without alignments)  
830.725 Million cell updates/sec

Title: US-10-089-452-24

Perfect score: 15

Sequence: 1 gacacctatgtgcac 15

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	14	93.3	185	4	US-09-382-552-38
C 2	14	93.3	537	4	US-10-101-464A-299
C 3	14	93.3	848	4	US-10-101-464A-298
C 4	14	93.3	2447	4	US-10-101-464A-871
C 5	14	93.3	13615	4	US-09-949-016-15643
C 6	13.4	89.3	15	2	US-08-232-081B-27
C 7	13.4	89.3	24	2	US-08-232-081B-18
C 8	13.4	89.3	25	4	US-09-396-196G-120502
C 9	13.4	89.3	25	4	US-09-396-196G-120503
C 10	13.4	89.3	25	4	US-09-396-196G-120504
C 11	13.4	89.3	54	4	US-08-454-899G-18
C 12	13.4	89.3	123	4	US-09-513-999C-29438
C 13	13.4	89.3	339	1	US-08-207-169A-1
C 14	13.4	89.3	354	2	US-08-232-081B-34
C 15	13.4	89.3	354	2	US-08-232-081B-37
C 16	13.4	89.3	360	1	US-08-463-128-1
C 17	13.4	89.3	360	1	US-08-463-298-1
C 18	13.4	89.3	360	2	US-08-436-339A-1
C 19	13.4	89.3	360	2	US-08-950-660-1
C 20	13.4	89.3	360	3	US-08-871-488A-14
C 21	13.4	89.3	360	4	US-08-454-899G-5
C 22	13.4	89.3	360	5	PCT-US93-00030-1
C 23	13.4	89.3	360	5	PCT-US93-00924-1
C 24	13.4	89.3	360	5	PCT-US94-10395-1
C 25	13.4	89.3	363	2	US-08-822-830B-1
C 26	13.4	89.3	363	2	US-08-822-830B-12
C 27	13.4	89.3	363	4	US-09-157-452B-1

28 13.4 89.3 363 4 US-09-157-452B-15 Sequence 15, Appl  
29 13.4 89.3 372 4 US-08-454-899G-46 Sequence 46, Appl  
30 13.4 89.3 409 1 US-08-024-253-1 Sequence 1, Appl  
31 13.4 89.3 409 4 US-09-513-999C-35635 Sequence 35635, A  
32 13.4 89.3 412 3 US-08-836-561-30 Sequence 30, Appl  
33 13.4 89.3 412 4 US-09-434-122-30 Sequence 30, Appl  
34 13.4 89.3 429 1 US-08-463-128-3 Sequence 3, Appl  
35 13.4 89.3 429 1 US-08-463-128-7 Sequence 7, Appl  
36 13.4 89.3 429 1 US-08-463-298-3 Sequence 3, Appl  
37 13.4 89.3 429 1 US-08-463-298-7 Sequence 3, Appl  
38 13.4 89.3 429 2 US-08-436-339A-3 Sequence 3, Appl  
39 13.4 89.3 429 3 US-08-436-339A-7 Sequence 28, Appl  
40 13.4 89.3 429 4 US-08-454-899G-28 Sequence 28, Appl  
41 13.4 89.3 429 4 US-08-454-899G-32 Sequence 32, Appl  
42 13.4 89.3 429 4 US-08-454-899G-38 Sequence 38, Appl  
43 13.4 89.3 429 4 US-08-454-899G-42 Sequence 42, Appl  
44 13.4 89.3 429 4 US-08-454-899G-50 Sequence 50, Appl  
45 13.4 89.3 429 4 US-08-454-899G-54 Sequence 54, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-382-552-38/c

; Sequence 38, Application US/09382552

; Patent No. 6673909

; GENERAL INFORMATION:

; APPLICANT: Brown, Jr., Robert H.

; APPLICANT: Liu, Jing

; APPLICANT: Aoki, Masashi

; APPLICANT: Ho, Meng

; APPLICANT: Maecuda-Asada, Chie

; TITLE OF INVENTION: DYSERLIN, A GENE MUTATED IN DISTAL MYOPATHY AND LIMB

; FILE REFERENCE: 00786/399002

; CURRENT APPLICATION NUMBER: US/09/382,552

; CURRENT FILING DATE: 1999-08-25

; EARLIER APPLICATION NUMBER: US 60/097,927

; EARLIER FILING DATE: 1998-08-25

; NUMBER OF SEQ ID NOS: 233

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 38

; LENGTH: 185

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-382-552-38

Query Match 93.3% Score 14; DB 4; Length 185;

Best Local Similarity 100.0%; Pred. No. 1.3e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14

Db 176 GACACCTATGTGCA 163

##### RESULT 2

US-10-101-464A-299/c

; Sequence 299, Application US/10101464A

; Patent No. 6768041

; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Nicolaas

; APPLICANT: Higgins, Colleen M.

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; FILE REFERENCE: 11000.1020C2 and Their Use in the Modification of Plant Cell Signaling

; CURRENT APPLICATION NUMBER: US/10/101,464A

; CURRENT FILING DATE: 2002-03-18

; PRIOR APPLICATION NUMBER: 09/704,302

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 09/228,986

100



Best Local Similarity 93.3%; Pred. No. 2.5e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
| | | | | | | | | |  
Db 5 GACACCTATATGCAC 19

## RESULT 10

US-09-396-196G-120504  
; Sequence 120504, Application US/09396196G  
; Patent No. 6821724  
; GENERAL INFORMATION:  
; APPLICANT: Michael Mittmann  
; APPLICANT: David Mack  
; APPLICANT: David Lockhart  
; APPLICANT: Affymetrix, Inc.  
; TITLE OF INVENTION: Methods of Genetic Analysis  
; FILE REFERENCE: 3101.1  
; CURRENT APPLICATION NUMBER: US/09/396.196G  
; CURRENT FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: 60/100,678  
; PRIOR FILING DATE: 1998-09-17  
; NUMBER OF SEQ ID NOS: 127806  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 120504  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: mus musculus  
US-09-396-196G-120504

Query Match 89.3%; Score 13.4; DB 4; Length 25;  
Best Local Similarity 93.3%; Pred. No. 2.5e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
| | | | | | | | | |  
Db 2 GACACCTATATGCAC 16

## RESULT 11

US-08-454-899G-18/c  
; Sequence 18, Application US/08454899G  
; Patent No. 6602503  
; GENERAL INFORMATION:  
; APPLICANT: Lobb, Roy R.  
; APPLICANT: Carr, Frank J.  
; APPLICANT: Tempest, Philip R.  
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES  
; FILE REFERENCE: 10274-007001  
; CURRENT APPLICATION NUMBER: US/08/454.899G  
; CURRENT FILING DATE: 1995-05-31  
; PRIOR APPLICATION NUMBER: US 08/004,798  
; PRIOR FILING DATE: 1993-01-12  
; PRIOR APPLICATION NUMBER: PCT/US94/00266  
; PRIOR FILING DATE: 1994-01-07  
; NUMBER OF SEQ ID NOS: 109  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 54  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: mutagenizing oligonucleotide  
US-08-454-899G-18

Query Match 89.3%; Score 13.4; DB 4; Length 54;  
Best Local Similarity 93.3%; Pred. No. 2.6e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
| | | | | | | | | |  
Db 26 GACACCTATATGCAC 12

## RESULT 12

US-09-513-999C-29438/c  
; Sequence 29438, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513.999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 29438  
; LENGTH: 123  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-513-999C-29438

Query Match 89.3%; Score 13.4; DB 4; Length 123;  
Best Local Similarity 93.3%; Pred. No. 2.8e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
| | | | | | | | | |  
Db 15 GACACCTATGTTCAC 1

## RESULT 13

US-08-207-169A-1  
; Sequence 1, Application US/08207169A  
; Patent No. 5674712  
; GENERAL INFORMATION:  
; APPLICANT: GRANDI, GUIDO  
; APPLICANT: DE PERA, FRANCESCA  
; APPLICANT: TOSI, CLAUDIO  
; APPLICANT: TORTORA, ORNELLA  
; APPLICANT: CUZZONI, ANNA  
; TITLE OF INVENTION: RECOMBINANT VECTOR AND USE THEREOF FOR  
; TITLE OF INVENTION: EXOCELLULAR PREPARATION OF SINGLE MOLECULA ANTIBODIES FROM  
; TITLE OF INVENTION: BACILLUS SUBTILIS  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/207,169A  
; FILING DATE: 08-MAR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5674712man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 2264-061-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR



```

; INFORMATION FOR SEQ ID NO: 1:
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 339 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;     FRAGMENT TYPE: N-terminal
;     ORIGINAL SOURCE:
;     CELL TYPE: Hybridoma
;     CELL LINE: 5E8
;     FEATURE:
;         NAME/KEY: CDS
;         LOCATION: 1..339
; US-08-2007-169A-1

```

Query Match 89.3%; Score 13.4; DB 1; Length 339;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels

```

RESULT 14
US-08-232-081B-34
; Sequence 34, Application US/08232081B
; Patent No. 5886152
; GENERAL INFORMATION:
; APPLICANT: NAKATANI, TOMOYUKI
; APPLICANT: GOMI, HIDEYUKI
; APPLICANT: WIJDENES, JOHN
; APPLICANT: NOGUCHI, HIROSHI
; TITLE OF INVENTION: HUMANIZED B-B10
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESSES:
; ADDRESSSES: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.081B

```

Query Match 89.3%; Score 13.4; DB 2; Length 354;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels

```

Db          91 GACACCTATATGAC 105
|||||
RESULT 15
US-08-232-081B-37
; Sequence #37, Application US/08232081B
; Patent No. 5886152
; GENERAL INFORMATION:
; APPLICANT: NAKATANI, TOMOYUKI
; APPLICANT: GOMI, HIDEYUKI
; APPLICANT: WIJENES, JOHN
; APPLICANT: NOGUCHI, HIROSHI
; TITLE OF INVENTION: HUMANIZED B-B10
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCI
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,081
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 20-3484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 354 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..354
US-08-232-081B-37

```

Query Match	89.3%	Score 13.4;	DB 2;	Length 354;
Best Local Similarity	93.3%;	Pred. No. 3.1e+02;		
Matches 14;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

RESULT 16  
US-08-463-128-1  
; Sequence 1, Application US/08463128  
; Patent No. 5695755  
; GENERAL INFORMATION:  
; APPLICANT: Papayannopoulou, Thalia (USA only)  
; APPLICANT: Board of Regents, U.  
; APPLICANT: Washington (except USA)  
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM  
; TITLE OF INVENTION: CELLS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: c/o FISH & NEAVE  
; STREET: 1251 Avenue of the Americas

```
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,128
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-11-1993
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-463-128-1

Query Match      89.3%; Score 13.4; DB 1; Length 360;
Best Local Similarity 93.3%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15
Db 88 GACACCTATATGCAC 102

RESULT 17
US-08-463-298-1
; Sequence 1, Application US/08463298
; Patent No. 5824304
; GENERAL INFORMATION:
; APPLICANT: Papayannopoulou, Thalia (USA only)
; APPLICANT: Board of Regents, U.
; APPLICANT: Washington (except USA)
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,298
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-NOV-1993
; APPLICATION NUMBER: US 07/977,702
```

```
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-463-298-1

Query Match      89.3%; Score 13.4; DB 1; Length 360;
Best Local Similarity 93.3%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15
Db 88 GACACCTATATGCAC 102

RESULT 18
US-08-436-339A-1
; Sequence 1, Application US/08436339A
; Patent No. 5843438
; GENERAL INFORMATION:
; APPLICANT: Papayannopoulou, Thalia (USA only)
; APPLICANT: Board of Regents, U.
; APPLICANT: Washington (except USA)
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,339A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,128
; FILING DATE: 05-JUN-1995
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-11-1993
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-436-339A-1
```

Query Match 89.3%; Score 13.4; DB 2; Length 360;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||||  
Db 88 GACACCTATGTGCAC 102

RESULT 19  
US-08-950-660-1  
; Sequence 1, Application US/08950660  
; Patent No. 5932214  
; GENERAL INFORMATION:  
; APPLICANT: Lobbs, Roy R.; Burkly, Linda C.  
; TITLE OF INVENTION: Treatment For Inflammatory Bowel Disease  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/950.660  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/456,124  
; FILING DATE:  
; APPLICATION NUMBER: US 08/373,857  
; FILING DATE: 18-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/284,603  
; FILING DATE: 11-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/00924  
; FILING DATE: 02-FEB-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/835,139  
; FILING DATE: 12-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Louis (PLM)  
; REGISTRATION NUMBER: 35,965-  
; REFERENCE/DOCKET NUMBER: BGP-031USCP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 360 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..360  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1  
; OTHER INFORMATION: /note= "PBAG159 insert: HPI/2 heavy  
; CHAIN INFORMATION: chain variable region; amino acid 1 is Glu (E) but  
; OTHER INFORMATION: Gln (Q) may be substituted"  
US-08-950-660-1

Query Match 89.3%; Score 13.4; DB 2; Length 360;

Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||||  
Db 88 GACACCTATGTGCAC 102

RESULT 20  
US-08-871-488A-14  
; Sequence 14, Application US/08871488A  
; Patent No. 6358710  
; GENERAL INFORMATION:  
; APPLICANT: Graves, Scott S.  
; APPLICANT: Reno, John M.  
; APPLICANT: Mallett, Robert W.  
; APPLICANT: Hylarides, Mark D.  
; APPLICANT: Searle, Stephen M.J.  
; APPLICANT: Henry, Andrew H.  
; APPLICANT: Pedersen, Jan T.  
; APPLICANT: Rees, Anthony R.  
; TITLE OF INVENTION: HUMANIZED ANTIBODIES THAT BIND TO THE  
; TITLE OF INVENTION: ANTIGEN BOUND BY ANTIBODY NR-LU-13 AND THEIR USE IN  
; TITLE OF INVENTION: PRETARGETING METHODS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/871.488A  
; FILING DATE: 09-JUN-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey Ph.D., Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 690022.527C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 360 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..360  
US-08-871-488A-14

Query Match 89.3%; Score 13.4; DB 3; Length 360;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||||  
Db 91 GACACCTATGTGCAC 105

RESULT 21  
US-08-454-899G-5  
; Sequence 5, Application US/08454899G  
; Patent No. 6602503  
; GENERAL INFORMATION:

```

RESULT 22
PCT-US93-00030-1
; Sequence 1, Application PC/TUS9300030
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; TITLE OF INVENTION: Treatment for Asthma
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegratti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00030
; FILING DATE: 19930112
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 92,307-A; D002 CIP PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1

```

```

Query Match      89.3%; Score 13.4; DB 5; Length 360;
Best Local Similarity 93.3%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCAC 15
   |||||
Db 88 GACACCTATGTGCAC 102

```



STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/822,830B  
FILING DATE: 03-MAR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/456,193  
FILING DATE: 31-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/374,331  
FILING DATE: 18-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/256,631  
FILING DATE: 12-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00030  
FILING DATE: 12-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/821,768  
FILING DATE: 13-JAN 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Myers, Louis (PLM)  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: BGP-021USCN  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 363 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..363  
NAME/KEY: misc\_feature  
LOCATION: 1  
OTHER INFORMATION: /note= "pBAG159 insert: HP1/2 heavy chain variable region"  
US-08-822-830B-12

Query Match 89.3%; Score 13.4; DB 2; Length 363;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
Db 91 GACACCTATATGCAC 105

RESULT 27  
US-09-157-452B-1  
Sequence 15, Application US/09157452B  
Patent No. 6482409  
GENERAL INFORMATION:  
APPLICANT: Lobb, Roy R.  
TITLE OF INVENTION: TREATMENT FOR INFLAMMATORY BOWEL DISEASE  
FILE REFERENCE: 10274-004003  
CURRENT APPLICATION NUMBER: US/09/157,452B  
CURRENT FILING DATE: 1998-09-21  
PRIOR APPLICATION NUMBER: US 08/950,660  
PRIOR FILING DATE: 1997-10-15  
PRIOR APPLICATION NUMBER: US 08/373,857

PRIOR FILING DATE: 1995-01-18  
PRIOR APPLICATION NUMBER: US 08/284,603  
PRIOR FILING DATE: 1994-08-11  
PRIOR APPLICATION NUMBER: PCT/US93/00924  
PRIOR FILING DATE: 1993-02-02  
PRIOR APPLICATION NUMBER: US 07/835,139  
PRIOR FILING DATE: 1992-02-12  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 363  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(363)  
NAME/KEY: misc\_feature  
LOCATION: (1)  
OTHER INFORMATION: /note= "pBAG159 insert: HP1/2 heavy chain variable regions; amino acid 1 is Glu (E) but Gln (Q) may be substituted as shown in US-09-157-452B-1"  
US-09-157-452B-1

Query Match 89.3%; Score 13.4; DB 4; Length 363;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
Db 91 GACACCTATATGCAC 105

RESULT 28  
US-09-157-452B-15  
Sequence 15, Application US/09157452B  
Patent No. 6482409  
GENERAL INFORMATION:  
APPLICANT: Lobb, Roy R.  
TITLE OF INVENTION: TREATMENT FOR INFLAMMATORY BOWEL DISEASE  
FILE REFERENCE: 10274-004003  
CURRENT APPLICATION NUMBER: US/09/157,452B  
CURRENT FILING DATE: 1998-09-21  
PRIOR APPLICATION NUMBER: US 08/950,660  
PRIOR FILING DATE: 1997-10-15  
PRIOR APPLICATION NUMBER: US 08/373,857  
PRIOR FILING DATE: 1995-01-18  
PRIOR APPLICATION NUMBER: US 08/284,603  
PRIOR FILING DATE: 1994-08-11  
PRIOR APPLICATION NUMBER: PCT/US93/00924  
PRIOR FILING DATE: 1993-02-02  
PRIOR APPLICATION NUMBER: US 07/835,139  
PRIOR FILING DATE: 1992-02-12  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 15  
LENGTH: 363  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(363)  
NAME/KEY: misc\_feature  
LOCATION: (1)  
OTHER INFORMATION: /note= "pBAG159 insert: HP1/2 heavy chain variable regions"  
US-09-157-452B-15

Query Match 89.3%; Score 13.4; DB 4; Length 363;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
Db 91 GACACCTATATGCAC 105

```
Db          91 GACACCTATATGCAC 105

RESULT 29
US-08-454-899G-46
; Sequence 46, Application US/08454899G
; Patent No. 6602503
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/08/454,899G
; CURRENT FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: heavy chain variable region
; NAME/KEY: CDS
; LOCATION: (1)...(372)
US-08-454-899G-46

Query Match      89.3%; Score 13.4; DB 4; Length 372;
Best Local Similarity 93.3%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          1 GACACCTATGTCAC 15
|||||
Db          91 GACACCTATATGCAC 105

RESULT 30
US-08-024-253-1
; Sequence 1, Application US/08024253
; Patent No. 5785968
; GENERAL INFORMATION:
; APPLICANT: KIMACHI, Kazuhiko
; APPLICANT: MAEDA, Hiroaki
; APPLICANT: NISHIYAMA, Kiyoto
; APPLICANT: TOKIYOSHI, Sachio
; APPLICANT: TOHIA, Yukinobu
; APPLICANT: MIKAMI, Takeshi
; TITLE OF INVENTION: ANTI-PELVINE CALCIVIRUS RECOMBINANT
; TITLE OF INVENTION: ANTIBODY AND GENE FRAGMENT ENCODING THE SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WEGNER, CANTOR, MUELLER & PLAYER
; STREET: 1233 20th Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-8218
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/024,253
; FILING DATE: 19930301
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 79189/1992

; FILING DATE: 28-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTOR, Herbert I.
; REGISTRATION NUMBER: 24,392
; REFERENCE/DOCKET NUMBER: P-500-23744
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-0400
; TELEFAX: (202) 835-0605
; TELEX: 440706 WEGHR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 409 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..408
US-08-024-253-1

Query Match      89.3%; Score 13.4; DB 1; Length 409;
Best Local Similarity 93.3%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy          1 GACACCTATGTCAC 15
|||||
Db          148 GACACCTATATGCAC 162

RESULT 31
US-09-513-999C-35635
; Sequence 35635, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 35635
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 137_
; OTHER INFORMATION: r-a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 139_
; OTHER INFORMATION: r-a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 146_
; OTHER INFORMATION: v-a or c or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 362_
; OTHER INFORMATION: v-a or c or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 370_
; OTHER INFORMATION: r-a or g
; FEATURE:
; NAME/KEY: misc_feature
```

; LOCATION: 385  
; OTHER INFORMATION: m=a or c  
US-09-513-999C-35635

Query Match 89.3%; Score 13.4; DB 4; Length 409;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 269 GACACCTATGTGCAC 283

## RESULT 32

US-08-836-561-30  
; Sequence 30, Application US/08836561  
; Patent No. 6018032

## GENERAL INFORMATION:

; APPLICANT: KOIKE, Masamichi  
; APPLICANT: FURUYA, Akiko  
; APPLICANT: NAKAMURA, Kazuyasu  
; APPLICANT: IIDA, Akihiro  
; APPLICANT: ANAZAWA, Hideharu  
; APPLICANT: HANAI, No. 6018032uo  
; APPLICANT: TAKATSU, Kiyoshi

; TITLE OF INVENTION: Antibody Against Human Interleukin-5  
; TITLE OF INVENTION: Receptor Alpha Chain

; NUMBER OF SEQUENCES: 106

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/836,561

; FILING DATE: 09-MAY-1997

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 232384/95

; FILING DATE: 11-SEP-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Lawrence, III, Stanton T

; REGISTRATION NUMBER: 25,736

; REFERENCE/DOCKET NUMBER: 7005-115-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090

; TELEFAX: 212-869-9741

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 30:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 412 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-08-836-561-30

Query Match 89.3%; Score 13.4; DB 3; Length 412;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 148 GACACCTATATGCAC 162

## RESULT 33

US-09-434-122-30

; Sequence 30, Application US/09434122

; Patent No. 6538111

; GENERAL INFORMATION:

; APPLICANT: KOIKE, Masamichi

; APPLICANT: FURUYA, Akiko

; APPLICANT: NAKAMURA, Kazuyasu

; APPLICANT: IIDA, Akihiro

; APPLICANT: ANAZAWA, Hideharu

; APPLICANT: HANAI, No. 6538111uo

; APPLICANT: TAKATSU, Kiyoshi

; TITLE OF INVENTION: Antibody Against Human Interleukin-5  
; TITLE OF INVENTION: Receptor Alpha Chain

; NUMBER OF SEQUENCES: 106

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/434,122

; FILING DATE: 05-NO. 6538111-1999

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/836,561

; FILING DATE: 09-MAY-1997

; APPLICATION NUMBER: JP 232384/95

; FILING DATE: 11-SEP-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Lawrence, III, Stanton T

; REGISTRATION NUMBER: 25,736

; REFERENCE/DOCKET NUMBER: 7005-115-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090

; TELEFAX: 212-869-9741

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 30:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 412 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 30:

US-09-434-122-30

Query Match 89.3%; Score 13.4; DB 4; Length 412;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 148 GACACCTATATGCAC 162

## RESULT 34

US-08-463-128-3

; Sequence 3, Application US/08463128

; Patent No. 5695755

; GENERAL INFORMATION:

; APPLICANT: Papayannopoulou, Thalia (USA only)

; APPLICANT: Board of Regents, U.

; APPLICANT: Washington (except USA)

; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM

; TITLE OF INVENTION: CELLS

; NUMBER OF SEQUENCES: 10



CORRESPONDENCE ADDRESS:  
ADDRESSEE: c/o FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,128  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/11060  
FILING DATE: 11-11-1993  
APPLICATION NUMBER: US 07/977,702  
FILING DATE: 11-11-1993  
APPLICATION NUMBER: US 07/977,702  
FILING DATE: 13-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B173CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9090  
TELEFAX: (212) 596-9090  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 429 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 1..57  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 58..429  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..429  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1  
OTHER INFORMATION: /note= "pMDR1019 insert: Stage 1  
OTHER INFORMATION: heavy chain variable region"  
US-08-463-128-3

Query Match 89.3%; Score 13.4; DB 1; Length 429;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTCAC 15  
|||||  
Db 148 GACACCTATGTCAC 162

RESULT 35  
US-08-463-128-7  
Sequence 7, Application US/08463128  
Patent No. 5695755  
GENERAL INFORMATION:  
APPLICANT: Papayannopoulou, Thalia (USA only)  
APPLICANT: Board of Regents, U.  
APPLICANT: Washington (except USA)  
TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM  
CELLS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: c/o FISH & NEAVE  
STREET: 1251 Avenue of the Americas

CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,128  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/11060  
FILING DATE: 11-11-1993  
APPLICATION NUMBER: US 07/977,702  
FILING DATE: 13-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr., James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: B173CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9090  
TELEFAX: (212) 596-9090  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 429 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 1..57  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 58..429  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..429  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1  
OTHER INFORMATION: /note= "pBAG195 insert: AS heavy  
OTHER INFORMATION: chain variable region"  
US-08-463-128-7

Query Match 89.3%; Score 13.4; DB 1; Length 429;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTCAC 15  
|||||  
Db 148 GACACCTATGTCAC 162

RESULT 36  
US-08-463-298-3  
Sequence 3, Application US/08463298  
Patent No. 5824304  
GENERAL INFORMATION:  
APPLICANT: Papayannopoulou, Thalia (USA only)  
APPLICANT: Board of Regents, U.  
APPLICANT: Washington (except USA)  
TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM  
CELLS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: c/o FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.

```
;
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,298
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-NOV-1993
; CLASSIFICATION: 424
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; NAME: Haley Jr., James F.
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173CIP
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 429 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 1..57
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 58..429
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..429
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: /note= "pMDR1019 insert: Stage 1
; OTHER INFORMATION: heavy chain variable region"
;
US-08-463-298-3

Query Match 89.3%; Score 13.4; DB 1; Length 429;
Best Local Similarity 93.3%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15
Db 148 GACACCTATATGCAC 162

RESULT 37
US-08-463-298-7
; Sequence 7, Application US/08463298
; Patent No. 5824304
; GENERAL INFORMATION:
; APPLICANT: Papayannopoulou, Thalia (USA only)
; APPLICANT: Board of Regents, U.
; APPLICANT: Washington (except USA)
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,298
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-NOV-1993
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; NAME: Haley Jr., James F.
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173CIP
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 429 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 1..57
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 58..429
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..429
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: /note= "pBAG195 insert: AS heavy
; OTHER INFORMATION: chain variable region"
;
US-08-463-298-7

Query Match 89.3%; Score 13.4; DB 1; Length 429;
Best Local Similarity 93.3%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15
Db 148 GACACCTATATGCAC 162

RESULT 38
US-08-436-339A-3
; Sequence 3, Application US/08436339A
; Patent No. 5843438
; GENERAL INFORMATION:
; APPLICANT: Papayannopoulou, Thalia (USA only)
; APPLICANT: Board of Regents, U.
; APPLICANT: Washington (except USA)
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,339A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,128
; FILING DATE: 05-JUN-1995
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-11-1993
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 429 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 1..57
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 58..429
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..429
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: /note= "pMDR1019 insert: Stage 1
; OTHER INFORMATION: heavy chain variable region"
; US-08-436-339A-3

Query Match 89.3%; Score 13.4; DB 2; Length 429;
Best Local Similarity 93.3%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15
Db 148 GACACCTATATGCAC 162

RESULT 39
US-08-436-339A-7
; Sequence 7, Application US/08436339A
; Patent No. 5843438
; GENERAL INFORMATION:
; APPLICANT: Papayannopoulou, Thalia (USA only)
; APPLICANT: Board of Regents, U.
; APPLICANT: Washington (except USA)
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/436,339A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,128
; FILING DATE: 05-JUN-1995
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-11-1993
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 429 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 1..57
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 58..429
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..429
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: /note= "pBAG195 insert: AS heavy
; OTHER INFORMATION: chain variable region"
; US-08-436-339A-7

Query Match 89.3%; Score 13.4; DB 2; Length 429;
Best Local Similarity 93.3%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15
Db 148 GACACCTATATGCAC 162

RESULT 40
US-08-454-899G-28
; Sequence 28, Application US/08454899G
; Patent No. 6602503
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/08/454,899G
; CURRENT FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: heavy chain variable region
; FEATURE:
```

; NAME/KEY: CDS  
; LOCATION: (1)...(429)  
US-08-454-899G-28

Query Match 89.3%; Score 13.4; DB 4; Length 429;  
Best Local Similarity 93.3%; Pred. No. 3.1e+02;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
| | | | | | | | | |  
Db 148 GACACCTATATGCAC 162

Search completed: June 4, 2005, 15:59:09  
Job time : 30.5455 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 14:31:52 ; Search time 118.636 Seconds  
(without alignments)  
777.260 Million cell updates/sec

Title: US-10-089-452-24

Perfect score: 15

Sequence: 1 gacacctatgtgcac 15

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	53809	17	US-10-292-798-1253
2	14	93.3	25	19	US-10-719-900-445854
3	14	93.3	185	10	US-09-382-860-38
4	14	93.3	247	9	US-09-998-598-2336
5	14	93.3	498	9	US-09-880-107-2559
6	14	93.3	528	13	US-10-027-632-20462
7	14	93.3	528	17	US-10-027-632-20462
8	14	93.3	537	14	US-10-101-464A-299
9	14	93.3	537	19	US-10-864-252-299
10	14	93.3	587	13	US-10-027-632-113123
11	14	93.3	587	13	US-10-027-632-113124

c 12	14	93.3	587	17	US-10-027-632-113123	Sequence 113123,
c 13	14	93.3	587	17	US-10-027-632-113124	Sequence 113124,
c 14	14	93.3	848	14	US-10-101-464A-298	Sequence 298, App
c 15	14	93.3	848	19	US-10-864-252-298	Sequence 298, App
c 16	14	93.3	919	17	US-10-425-114-10392	Sequence 10392, A
c 17	14	93.3	1078	17	US-10-424-599-39503	Sequence 39503, A
c 18	14	93.3	1080	17	US-10-425-114-17594	Sequence 17594, A
c 19	14	93.3	1089	18	US-10-425-115-44423	Sequence 44423, A
c 20	14	93.3	1394	18	US-10-739-930-2659	Sequence 2659, Ap
c 21	14	93.3	2447	14	US-10-101-464A-871	Sequence 871, App
c 22	14	93.3	2447	19	US-10-864-252-871	Sequence 871, App
c 23	14	93.3	2758	17	US-10-094-749-595	Sequence 595, App
c 24	14	93.3	2843	17	US-10-108-260A-1498	Sequence 1498, Ap
c 25	14	93.3	5092	17	US-10-621-758A-43	Sequence 43, Appl
c 26	14	93.3	5092	18	US-10-863-208A-43	Sequence 43, Appl
c 27	14	93.3	5092	18	US-10-646-301A-43	Sequence 43, Appl
c 28	14	93.3	5092	18	US-10-736-769-43	Sequence 1, Appli
c 29	14	93.3	21501	17	US-10-377-315-1	Sequence 6831, Ap
c 30	14	93.3	80399	18	US-10-719-993-6831	Sequence 34, Appl
c 31	14	93.3	94752	17	US-10-408-168-34	Sequence 1, Appli
c 32	14	93.3	130320	17	US-10-408-168-1	Sequence 66, Appl
c 33	14	93.3	213040	13	US-10-087-192-856	Sequence 2, Appli
c 34	14	93.3	252907	18	US-10-417-375-66	Sequence 252811,
c 35	14	93.3	483728	18	US-10-699-156-2	Sequence 252811,
c 36	13.6	90.7	573	13	US-10-027-632-252811	Sequence 252811,
c 37	13.6	90.7	573	17	US-10-027-632-252811	Sequence 148539,
c 38	13.6	90.7	769	13	US-10-027-632-148539	Sequence 148540,
c 39	13.6	90.7	769	13	US-10-027-632-148540	Sequence 148541,
c 40	13.6	90.7	769	13	US-10-027-632-148541	Sequence 148539,
c 41	13.6	90.7	769	17	US-10-027-632-148539	Sequence 148540,
c 42	13.6	90.7	769	17	US-10-027-632-148540	Sequence 148541,
c 43	13.6	90.7	769	17	US-10-027-632-148541	Sequence 186602,
c 44	13.4	89.3	25	19	US-10-719-900-186602	Sequence 282774,
c 45	13.4	89.3	25	19	US-10-719-900-282774	

#### ALIGNMENTS

#### RESULT 1

US-10-292-798-1253

; Sequence 1253, Application US/10292798

; Publication No. US20030235833A1

; GENERAL INFORMATION:

; APPLICANT: SUWA, MAKIKO

; APPLICANT: ASAI, KIYOSHI

; APPLICANT: AKIYAMA, YUTAKA

; APPLICANT: ABURATANI, HIROYUKI

; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS

; FILE REFERENCE: 084335/166

; CURRENT APPLICATION NUMBER: US/10/292,798

; PRIOR FILING DATE: 2002-11-13

; PRIOR FILING DATE: 2002-11-13

; PRIOR FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: JP 2001-246789

; PRIOR FILING DATE: 2001-06-18

; NUMBER OF SEQ ID NOS: 2070

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1253

; LENGTH: 53809

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; LOCATION: source

; FEATURE:

; LOCATION: (1)..(53809)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (201)..(288)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1271)..(1350)

; FEATURE:

```
; NAME/KEY: CDS
; LOCATION: (1650)..(1677)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3021)..(3078)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8331)..(8353)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10809)..(10841)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12020)..(12064)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12426)..(12862)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13717)..(13755)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19514)..(19586)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25189)..(25368)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (27408)..(27479)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (27653)..(27826)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (28136)..(28300)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (30726)..(30896)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34508)..(34678)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (36623)..(36645)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37476)..(37617)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (38132)..(38331)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39960)..(40065)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (40143)..(40261)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (43119)..(43365)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (43640)..(43792)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (45218)..(45540)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (45967)..(46091)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48355)..(48520)
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (48693)..(48820)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (49156)..(49202)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51179)..(51256)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51673)..(52253)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53190)..(53609)
; US-10-292-798-1253
```

```
Query Match 100.0%; Score 15; DB 17; Length 53809;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GACACCTATGTGCAC 15
|||||
Db 4195 GACACCTATGTGCAC 4209
```

## RESULT 2

```
US-10-719-900-445854
; Sequence 445854, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 445854
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-719-900-445854
```

```
Query Match 93.3%; Score 14; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GACACCTATGTGCA 14
|||||
Db 1 GACACCTATGTGCA 14
```

## RESULT 3

```
US-09-382-860-38/c
; Sequence 38, Application US/09382860
; Publication No. US20030110526A1
; GENERAL INFORMATION:
; APPLICANT: Brown, Jr., Robert H.
; APPLICANT: Liu, Jing
; APPLICANT: Aoki, Masashi
; APPLICANT: Hoffman, Eric
; APPLICANT: Chou, Fan-Li
; TITLE OF INVENTION: DYSPERLIN MUTATIONS
; FILE REFERENCE: 00786/401002
; CURRENT APPLICATION NUMBER: US/09/382,860
; CURRENT FILING DATE: 1999-08-25
; EARLIER APPLICATION NUMBER: US 60/097,930
; EARLIER FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 185
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-382-860-38

Query Match      93.3%; Score 14; DB 10; Length 185;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14
Db 176 GACACCTATGTGCA 163

RESULT 4
US-09-998-598-2336
; Sequence 2336, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Meagher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 2336
; LENGTH: 247
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-2336

Query Match      93.3%; Score 14; DB 9; Length 247;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14
Db 117 GACACCTATGTGCA 130

RESULT 5
US-09-880-107-2559
; Sequence 2559, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR FILING DATE: 2001-06-14
; PRIOR FILING DATE: 2000-06-14
; PRIOR FILING DATE: 2000-06-14
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2559
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N53757
; NAME/KEY: unsure
; LOCATION: (1)..(498)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2559

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-382-860-38

Query Match      93.3%; Score 14; DB 10; Length 185;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14
Db 176 GACACCTATGTGCA 163

RESULT 4
US-09-998-598-2336
; Sequence 2336, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Meagher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 2336
; LENGTH: 247
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-2336

Query Match      93.3%; Score 14; DB 9; Length 247;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14
Db 117 GACACCTATGTGCA 130

RESULT 5
US-09-880-107-2559
; Sequence 2559, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR FILING DATE: 2001-06-14
; PRIOR FILING DATE: 2000-06-14
; PRIOR FILING DATE: 2000-06-14
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2559
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N53757
; NAME/KEY: unsure
; LOCATION: (1)..(498)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2559

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-382-860-38

Query Match      93.3%; Score 14; DB 9; Length 498;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACCTATGTGCAC 15
Db 36 ACACCTATGTGCAC 49

RESULT 6
US-10-027-632-20462
; Sequence 20462, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20462
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-20462

Query Match      93.3%; Score 14; DB 13; Length 528;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACCTATGTGCAC 15
Db 202 ACACCTATGTGCAC 215

RESULT 7
US-10-027-632-20462
; Sequence 20462, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
```

```

; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20462
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-20462

```

Query Match 93.3%; Score 14; DB 17; Length 528;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Query Match          93.3%; Score 14; DB 14; Length 537;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 299
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-864-252-299

```

Query Match 93.3%; Score 14; DB 19; Length 537;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match	93.3%	Score 14;	DB 13;	Length 587;
Best Local Similarity	100.0%;	Pred. NO. 3.2e-02;		
Matches 14;	Conservative 0;	Mismatches 0;	Indels 0;	Caps 0;



APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR FILING DATE: 2000-07-12 US 60/218,006  
; PRIOR FILING DATE: 2000-04-20 US 60/198,676  
; PRIOR FILING DATE: 2000-04-20 US 60/218,006  
; PRIOR FILING DATE: 2000-03-29 US 60/193,483  
; PRIOR FILING DATE: 2000-03-29 US 60/185,218  
; PRIOR FILING DATE: 2000-02-24 US 60/167,363  
; PRIOR FILING DATE: 1999-11-23 US 60/156,358  
; PRIOR FILING DATE: 1999-09-28 US 60/146,002  
; PRIOR FILING DATE: 1999-08-09 US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 113124  
; LENGTH: 587  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-113124

Query Match 93.3%; Score 14; DB 13; Length 587;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACCTATGTGCAC 15  
Db 140 ACACCTATGTGCAC 127

## RESULT 12

US-10-027-632-113123/c  
; Sequence 113123, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12 US 60/218,006  
; PRIOR FILING DATE: 2000-04-20 US 60/198,676  
; PRIOR FILING DATE: 2000-04-20 US 60/218,006  
; PRIOR FILING DATE: 2000-03-29 US 60/193,483  
; PRIOR FILING DATE: 2000-03-29 US 60/185,218  
; PRIOR FILING DATE: 2000-02-24 US 60/167,363  
; PRIOR FILING DATE: 1999-11-23 US 60/156,358  
; PRIOR FILING DATE: 1999-09-28 US 60/146,002  
; PRIOR FILING DATE: 1999-08-09 US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 113123  
; LENGTH: 587  
; TYPE: DNA  
; ORGANISM: Human

Query Match 93.3%; Score 14; DB 17; Length 587;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACCTATGTGCAC 15  
Db 140 ACACCTATGTGCAC 127

## RESULT 13

US-10-027-632-113124/c  
; Sequence 113124, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12 US 60/218,006  
; PRIOR FILING DATE: 2000-04-20 US 60/198,676  
; PRIOR FILING DATE: 2000-04-20 US 60/218,006  
; PRIOR FILING DATE: 2000-03-29 US 60/193,483  
; PRIOR FILING DATE: 2000-03-29 US 60/185,218  
; PRIOR FILING DATE: 2000-02-24 US 60/167,363  
; PRIOR FILING DATE: 1999-11-23 US 60/156,358  
; PRIOR FILING DATE: 1999-09-28 US 60/146,002  
; PRIOR FILING DATE: 1999-08-09 US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 113124  
; LENGTH: 587  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-113124

Query Match 93.3%; Score 14; DB 17; Length 587;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACCTATGTGCAC 15  
Db 140 ACACCTATGTGCAC 127

## RESULT 14

US-10-101-464A-298/c  
; Sequence 298, Application US/10101464A  
; Publication No. US20030046728A1  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Higgins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; FILE REFERENCE: 11000.1020C2  
; CURRENT APPLICATION NUMBER: US/10/101,464A  
; CURRENT FILING DATE: 2002-03-18  
; PRIOR APPLICATION NUMBER: 09/704,302  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 09/228,986  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/162,866  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT/US00/00724  
; PRIOR FILING DATE: 2000-01-11  
; NUMBER OF SEQ ID NOS: 989  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 298  
; LENGTH: 848

; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-10-101-464A-298

Query Match 93.3%; Score 14; DB 14; Length 848;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15  
| | | | | | | | | | | | | | | |  
Db 142 ACACCTATGTGCAC 129

## RESULT 15

US-10-864-252-298/c  
; Sequence 298, Application US/10864252  
; Publication No. US20050050583A1  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Higgins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; FILE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
; FILE REFERENCE: 11000.1020c3  
; CURRENT APPLICATION NUMBER: US/10/864,252  
; CURRENT FILING DATE: 2004-06-09  
; PRIOR APPLICATION NUMBER: 10/101,464  
; PRIOR FILING DATE: 2002-03-18  
; PRIOR APPLICATION NUMBER: 09/704,302  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 09/228,986  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/162,866  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT/US00/00724  
; PRIOR FILING DATE: 2000-01-11  
; NUMBER OF SEQ ID NOS: 989  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 298  
; LENGTH: 848  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-10-864-252-298

Query Match 93.3%; Score 14; DB 19; Length 848;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15  
| | | | | | | | | | | | | | | |  
Db 142 ACACCTATGTGCAC 129

## RESULT 16

US-10-425-114-10392  
; Sequence 10392, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 10392  
; LENGTH: 919  
; TYPE: DNA

; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700907070\_FLI  
US-10-425-114-10392

Query Match 93.3%; Score 14; DB 17; Length 919;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15  
| | | | | | | | | | | | | | | |  
Db 406 ACACCTATGTGCAC 419

## RESULT 17

US-10-424-599-39503  
; Sequence 39503, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 39503  
; LENGTH: 1078  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_135670C.1  
US-10-424-599-39503

Query Match 93.3%; Score 14; DB 17; Length 1078;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15  
| | | | | | | | | | | | | | | |  
Db 482 ACACCTATGTGCAC 495

## RESULT 18

US-10-425-114-17594/c  
; Sequence 17594, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 17594  
; LENGTH: 1080  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3069-052-E7\_FLI  
US-10-425-114-17594

Query Match 93.3%; Score 14; DB 17; Length 1080;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;

```
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACACCTATGTGCA 14
| | | | | | | | | |
Db 738 GACACCTATGTGCA 725

RESULT 19
US-10-425-115-44423/c
; Sequence 44423, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 44423
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_140519C.1
; US-10-425-115-44423

Query Match 93.3%; Score 14; DB 18; Length 1089;
Best Local Similarity 100.0%; Pred. No. 3.3e+02; Length 1089;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACACCTATGTGCA 14
| | | | | | | | | |
Db 738 GACACCTATGTGCA 725

RESULT 20
US-10-739-930-2659/c
; Sequence 2659, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 2659
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER4353_3
; US-10-739-930-2659

Query Match 93.3%; Score 14; DB 18; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3.3e+02; Length 1394;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACACCTATGTGCA 14
| | | | | | | | | |
Db 908 GACACCTATGTGCA 895

RESULT 21
US-10-101-464A-871/c
; Sequence 871, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 871
; LENGTH: 2447
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; US-10-101-464A-871

Query Match 93.3%; Score 14; DB 19; Length 2447;
Best Local Similarity 100.0%; Pred. No. 3.4e+02; Length 2447;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 ACACCTATGTGCAC 15
| | | | | | | | | |
Db 156 ACACCTATGTGCAC 143

RESULT 22
US-10-864-252-871/c
; Sequence 871, Application US/10864252
; Publication No. US20050050583A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c3
; CURRENT APPLICATION NUMBER: US/10/864,252
; CURRENT FILING DATE: 2004-06-09
; PRIOR APPLICATION NUMBER: 10/101,464
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 871
; LENGTH: 2447
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; US-10-864-252-871

Query Match 93.3%; Score 14; DB 19; Length 2447;
Best Local Similarity 100.0%; Pred. No. 3.4e+02; Length 2447;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 ACACCTATGTGCAC 15
| | | | | | | | | |
Db 156 ACACCTATGTGCAC 143
```

## RESULT 23

US-10-094-749-595/c  
; Sequence 595, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHIKO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOTYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 595  
; LENGTH: 2758  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-094-749-595

Query Match 93.3%; Score 14; DB 17; Length 2758;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14  
Db 1092 GACACCTATGTGCA 1079

## RESULT 24

US-10-108-260A-1498  
; Sequence 1498, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1e1 full length cdna  
; FILE REFERENCE: HL-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1498  
; LENGTH: 2843  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-108-260A-1498

Query Match 93.3%; Score 14; DB 17; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACCTATGTGCAC 15  
Db 1116 ACACCTATGTGCAC 1129

## RESULT 25

US-10-621-758A-43/c  
; Sequence 43, Application US/10621758A  
; Publication No. US20040093629A1  
; GENERAL INFORMATION:  
; APPLICANT: Altmann, Scott W  
; APPLICANT: Wang, Luquan  
; APPLICANT: Graziano, Michael  
; APPLICANT: Murgolo, Nick  
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF  
; FILE REFERENCE: JB01603-K-US  
; CURRENT APPLICATION NUMBER: US/10/621,758A  
; CURRENT FILING DATE: 2003-07-17  
; PRIOR APPLICATION NUMBER: 60/397,442  
; PRIOR FILING DATE: 2002-07-19  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 43  
; LENGTH: 5092  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (57)..(4136)  
; OTHER INFORMATION:  
US-10-621-758A-43

Query Match 93.3%; Score 14; DB 17; Length 5092;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACCTATGTGCAC 15  
Db 5024 ACACCTATGTGCAC 5011

## RESULT 26

US-10-663-208A-43/c  
; Sequence 43, Application US/10663208A  
; Publication No. US20040132058A1  
; GENERAL INFORMATION:  
; APPLICANT: Altmann, Scott W  
; APPLICANT: Wang, Luquan  
; APPLICANT: Graziano, Michael  
; APPLICANT: Murgolo, Nick  
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF  
; FILE REFERENCE: JB01603K2 US  
; CURRENT APPLICATION NUMBER: US/10/663,208A  
; CURRENT FILING DATE: 2003-09-16  
; PRIOR APPLICATION NUMBER: 60/397,442  
; PRIOR FILING DATE: 2002-07-19  
; PRIOR APPLICATION NUMBER: 10/621,758  
; PRIOR FILING DATE: 2003-07-17  
; PRIOR APPLICATION NUMBER: 10/646,301  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 43  
; LENGTH: 5092  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (57)..(4136)  
; OTHER INFORMATION:  
US-10-663-208A-43

Query Match 93.3%; Score 14; DB 18; Length 5092;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      2 ACACCTATGTGCAC 15
Db      5024 ACACCTATGTGCAC 5011

RESULT 27
US-10-646-301A-43/c
; Sequence 43, Application US/10646301A
; Publication No. US20040137467A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)..(4136)
; OTHER INFORMATION:
US-10-646-301A-43

Query Match      93.3%; Score 14; DB 18; Length 5092;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ACACCTATGTGCAC 15
Db      5024 ACACCTATGTGCAC 5011

RESULT 28
US-10-736-769-43/c
; Sequence 43, Application US/10736769
; Publication No. US20040161838A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-US
; CURRENT APPLICATION NUMBER: US/10/736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)..(4136)

US-10-736-769-43

Query Match      93.3%; Score 14; DB 18; Length 5092;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ACACCTATGTGCAC 15
Db      5024 ACACCTATGTGCAC 5011

RESULT 29
US-10-377-315-1/c
; Sequence 1, Application US/10377315
; Publication No. US20030229041A1
; GENERAL INFORMATION:
; APPLICANT: Sutherland, May S. Kung
; APPLICANT: Geoghegan, James Charles
; APPLICANT: Yu, Changpu
; APPLICANT: Latham, John
; APPLICANT: Celltech R&D, Inc.
; TITLE OF INVENTION: Methods to Increase or Decrease Bone Density
; FILE REFERENCE: 1427.005US1
; CURRENT APPLICATION NUMBER: US/10/377,315
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/361,258
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/406,171
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: US Ser. No. US20030229041A1 Unknown
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 21501
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-377-315-1

Query Match      93.3%; Score 14; DB 17; Length 21501;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACACCTATGTGCA 14
Db      2535 GACACCTATGTGCA 2522

RESULT 30
US-10-719-993-6831
; Sequence 6831, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; FILE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6831
; LENGTH: 80399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-6831

Query Match      93.3%; Score 14; DB 18; Length 80399;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ACACCTATGTGCAC 15
```

```
Db 45236 ACACCTATGTGCAC 45249
|||||
RESULT 31
US-10-408-168-34
; Sequence 34, Application US/10408168
; Publication No. US20030235847A1
; GENERAL INFORMATION:
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Proll, Sean
; APPLICANT: Charmley, Patrick R.
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Uitterlinden, Andreas Gerardus
; TITLE OF INVENTION: ASSOCIATION OF POLYMORPHISMS IN THE SOST
; TITLE OF INVENTION: GENE REGION WITH BONE MINERAL DENSITY
; FILE REFERENCE: 240083.525
; CURRENT APPLICATION NUMBER: US/10/408,168
; CURRENT FILING DATE: 2003-04-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 94752
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-408-168-34

Query Match 93.3%; Score 14; DB 17; Length 94752;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14
|||||
Db 18695 GACACCTATGTGCA 18708
|||||

RESULT 32
US-10-408-168-1/c
; Sequence 1, Application US/10408168
; Publication No. US20030235847A1
; GENERAL INFORMATION:
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Proll, Sean
; APPLICANT: Charmley, Patrick R.
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Uitterlinden, Andreas Gerardus
; TITLE OF INVENTION: ASSOCIATION OF POLYMORPHISMS IN THE SOST
; TITLE OF INVENTION: GENE REGION WITH BONE MINERAL DENSITY
; FILE REFERENCE: 240083.525
; CURRENT APPLICATION NUMBER: US/10/408,168
; CURRENT FILING DATE: 2003-04-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 130320
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 100722, 100754, 102080, 117731, 124408, 124532, 124585, 124955,
; LOCATION: 124956, 124963, 124964, 124965, 124967
; OTHER INFORMATION: n = A,T,C or G
US-10-408-168-1

Query Match 93.3%; Score 14; DB 17; Length 130320;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14
|||||
Db 2536 GACACCTATGTGCA 2523
|||||

RESULT 33
US-10-087-192-856
; Sequence 856, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 856
; LENGTH: 213040
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(213040)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-856

Query Match 93.3%; Score 14; DB 13; Length 213040;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14
|||||
Db 61402 GACACCTATGTGCA 61415
|||||

RESULT 34
US-10-417-375-66/c
; Sequence 66, Application US/10417375
; Publication No. US20040219528A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001600
; CURRENT APPLICATION NUMBER: US/10/417,375
; CURRENT FILING DATE: 2003-04-15
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 252907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-417-375-66

Query Match 93.3%; Score 14; DB 18; Length 252907;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14
|||||
Db 72693 GACACCTATGTGCA 72680
|||||

RESULT 35
US-10-699-156-2/c
; Sequence 2, Application US/10699156
; Publication No. US20040197799A1
; GENERAL INFORMATION:
; APPLICANT: Williamson, Robert
; APPLICANT: Dahl, Hans-Henrik
; APPLICANT: Forrest, Susan
```

```
; APPLICANT: Delatycki, Martin
; APPLICANT: Wilcox, Stephen
; APPLICANT: de Silva, Michelle
; APPLICANT: Elliott, Katherine
; APPLICANT: Lynch, Michael
; FILE OF INVENTION: DETERMINATION OF A GENETIC PREDISPOSITION FOR BEHAVIORAL DISORDER
; TITLE REFERENCE: A36055-PCT-USA-A 071838.0143
; CURRENT APPLICATION NUMBER: US/10/699,156
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: PCT/AU02/00556
; PRIOR APPLICATION NUMBER: 2002-05-03
; PRIOR APPLICATION NUMBER: AU PR4756
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 60/295811
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: AU PR5426
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 483728
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (19734)..(19962)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (69768)..(69974)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (94975)..(95181)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (127559)..(130531)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142395)..(143725)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152195)..(154039)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (167914)..(168120)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (170745)..(173326)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (179281)..(182018)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (208000)..(209952)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (224704)..(224953)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231493)..(233474)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (239274)..(247177)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (262462)..(263424)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (283171)..(283390)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (285365)..(288406)
; OTHER INFORMATION: n = any nucleotide
; US-10-699-156-2
```

```
Query Match 93.3%; Score 14; DB 18; Length 483728;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 ACACCTATGTGCAC 15
Db 7616 ACACCTATGTGCAC 7603
|||||
```

## RESULT 36

```
US-10-027-632-252811/c
; Sequence 252811, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252811
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-252811
```

```
Query Match 90.7%; Score 13.6; DB 13; Length 573;
Best Local Similarity 92.9%; Pred. No. 5.6e+02;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GACACCTATGTGCA 14
Db 417 GACACCTATGTGCA 404
|||||
```

## RESULT 37

```
US-10-027-632-252811/c
; Sequence 252811, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 252811
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-252811

Query Match          90.7%; Score 13.6; DB 17; Length 573;
Best Local Similarity 92.9%; Pred. No. 5.6e+02;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GACACCTATGTGCA 14
Db      417 GACACCTATGTGCA 404
|||||:|||||:|||||

RESULT 38
US-10-027-632-148539
; Sequence 148539, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 148539
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-148539

Query Match          90.7%; Score 13.6; DB 13; Length 769;
Best Local Similarity 92.9%; Pred. No. 5.6e+02;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 ACACCTATGTGCAC 15
Db      249 ACACMTATGTGCAC 262
|||||:|||||:|||||

RESULT 39
US-10-027-632-148540
; Sequence 148540, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 148540
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-148540

Query Match          90.7%; Score 13.6; DB 13; Length 769;
Best Local Similarity 92.9%; Pred. No. 5.6e+02;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 ACACCTATGTGCAC 15
Db      249 ACACMTATGTGCAC 262
|||||:|||||:|||||

RESULT 40
US-10-027-632-148541
; Sequence 148541, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 148539
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-148539

Query Match          90.7%; Score 13.6; DB 13; Length 769;
Best Local Similarity 92.9%; Pred. No. 5.6e+02;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2 ACACCTATGTGCAC 15
Db      249 ACACMTATGTGCAC 262
|||||:|||||:|||||

RESULT 39
US-10-027-632-148540
; Sequence 148540, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 148540
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-148540

Query Match          90.7%; Score 13.6; DB 13; Length 769;
Best Local Similarity 92.9%; Pred. No. 5.6e+02;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 ACACCTATGTGCAC 15
Db      249 ACACMTATGTGCAC 262
|||||:|||||:|||||

RESULT 40
US-10-027-632-148541
; Sequence 148541, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 148540
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-148540

Query Match          90.7%; Score 13.6; DB 13; Length 769;
Best Local Similarity 92.9%; Pred. No. 5.6e+02;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```



```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 148541
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-148541

Query Match      90.7%; Score 13.6; DB 13; Length 769;
Best Local Similarity 92.9%; Pred. No. 5.6e+02;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 ACACCTATGTGCAC 15
      ||||:|||||
Db      249 ACACWTATGTGCAC 262

Search completed: June 4, 2005, 18:27:19
Job time : 121.636 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 12:51:52 ; Search time 766.212 Seconds  
(without alignments)  
745.178 Million cell updates/sec

Title: US-10-089-452-24

Perfect score: 15

Sequence: 1 gacacatgtgcac 15

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_hic.\*

4: gb\_est3.\*

5: gb\_est4.\*

6: gb\_est5.\*

7: gb\_est6.\*

8: gb\_ges1.\*

9: gb\_ges2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	15	100.0	551	8	AQ970549	RPCI-23-3
2	15	100.0	685	6	CD482717	atr01-23m
3	15	100.0	698	8	AQ378906	RPCI11-16
4	15	100.0	711	4	BI086158	602870041
5	15	100.0	754	9	AG480637	Mus muscu
6	15	100.0	1030	4	BM322137	RM79773 S
7	14	93.3	62	8	BH908390	SAUK_0478
8	14	93.3	193	7	CK578709	IST_W15.1
9	14	93.3	255	6	CD340276	StrPus36.
10	14	93.3	273	1	AA558239	nl34905.s
11	14	93.3	289	1	AI702394	tz67b08.x
12	14	93.3	291	1	AA676945	zj69h02.s
13	14	93.3	293	1	AL601167	DFZp313B
14	14	93.3	316	6	CD599154	RK115A1H0
15	14	93.3	345	2	BF423663	er38b12.y
16	14	93.3	353	4	BJ695065	BJ695065
17	14	93.3	354	7	CK926149	p3fmgc.01
18	14	93.3	360	2	BB841157	BB841157
19	14	93.3	375	9	CL327998	RPCI14_24
20	14	93.3	381	2	BB817276	BB817276
21	14	93.3	383	6	BY670540	BY670540
22	14	93.3	391	6	BY615493	BY615493
23	14	93.3	397	8	AQ597571	HS_2030 A
24	14	93.3	399	1	AA632057	np70e01.s

C 25	14	93.3	402	2	AW354829	37769 MAR
C 26	14	93.3	403	6	BY624835	BY624835
C 27	14	93.3	404	8	AQ204278	HS_3113 B
C 28	14	93.3	406	7	H25220	Ym56R08.t1
C 29	14	93.3	409	6	CB770156	AMGNNUC:S
C 30	14	93.3	420	5	BQ296291	san89f05.
C 31	14	93.3	420	5	BY370371	BY370371
C 32	14	93.3	420	9	CG985509	CH240.155
C 33	14	93.3	421	7	N58741	Y74d12.t1
C 34	14	93.3	422	7	R53359	Y983e08.t1
C 35	14	93.3	423	7	CR555382	DKFZp459D
C 36	14	93.3	424	2	BB815593	BB815593
C 37	14	93.3	425	1	AU018082	AU018082
C 38	14	93.3	426	2	AW471444	xw59e10.x
C 39	14	93.3	427	5	BQ559024	H4055F08-
C 40	14	93.3	429	2	BE074759	IL5-BT057
C 41	14	93.3	432	2	AW099707	ed30b09.y
C 42	14	93.3	434	2	BE074758	IL5-BT057
C 43	14	93.3	434	7	CK337677	C0340A03-
C 44	14	93.3	436	2	BE931620	QV2-HT049
C 45	14	93.3	437	1	AI457611	tj63c12.x

#### ALIGNMENTS

RESULT 1

AQ970549

LOCUS

DEFINITION

RPCI-23-330D21.TVB RPCI-23 Mus musculus genomic clone

ACCESSION

AQ970549

VERSION

AQ970549.1 GI:6801002

KEYWORDS

GSS.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Jong, P. and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-23

TITLE

JOURNAL

COMMENT

Unpublished (1999)

Other\_GSSs: RPCI-23-330D21.TVB

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong

(pieter@jeong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)

or from Resea ch Genetics (info@resgen.com). BAC end page:

http://www.tigr.org/tdb/bac ends/mouse/bac\_end\_intro.html

Plate: 330 row: D column: 21

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1. 551

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="RPCI-23-330D21"

/sex="Female"

/lab host="DH10B"

/clone lib="RPCI-23"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site: 1;

ECORI; Site\_2: EcoRI; Female C57BL/6J mouse kidney\_and/or

brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRII. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies). "

## ORIGIN

Query Match 100.0%; Score 15; DB 8; Length 551;  
Best Local Similarity 100.0%; Pred. No. 9.5e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 273 GACACCTATGTGCAC 287

RESULT 2  
CD482717  
LOCUS  
DEFINITION  
CD482717 685 bp mRNA linear EST 04-JUN-2003  
atr01-23msl-c05 atr01 Amborella trichopoda cDNA clone  
CD482717  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Amborella trichopoda

CD482717.1 GI:31403985  
EST.  
Amborella trichopoda  
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; stem Magnoliophyta; Amborellaceae;  
Amborella.

REFERENCE  
AUTHORS  
1 (bases 1 to 685)  
Gepamphilis,C., Soltis,D., Soltis,P., Farmerie,W., Ma,H.,  
Tankley,S., Leebens-Mack,J., Field,D., Buzgo,M., Kim,S., Ilut,D.,  
Landherr,L., Hu,Y., Wall,K., Albert,V., Carlson,J., Doyle,J.,  
Frohlich,M., Miller,W., Oppenheimer,D. and Theissen,G.  
Generation of ESTs from early flower buds of Amborella trichopoda  
Unpublished (2002)  
JOURNAL  
COMMENT  
Contact: Claude dePamphilis or James Leebens-Mack  
Mueller Laboratory  
Penn State University  
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn  
State University, University Park, PA 16802, USA  
Tel: 814 863 6413  
Fax: 814 865 9131  
Email: cwd3@psu.edu or jhl10@psu.edu

The sequence provided is trimmed of vector and low quality regions.  
Full sequence and original trace file are available from the Plant  
Genome Network website (<http://pgn.cornell.edu>)  
Plate: atr01-23msl row: c column: 05  
Seq primer: M13F.

FEATURES  
source

Location/Qualifiers  
1..685  
/organism="Amborella trichopoda"  
/mol\_type="mRNA"  
/db\_xref="taxon:13333"  
/clone="atr01-23msl-c05"  
/tissue\_type="flower buds"  
/dev\_stage="<= 2.5mm buds"  
/lab\_host="SOLR"  
/clone\_lib="atr01"

/note="Vector: pBluescript SK (+/-); Site\_1: EcoRI;  
Site\_2: XhoI; Amborella trichopoda Baili; This library was  
made from male flowers only. Only floral buds with  
diameter of 2.5 mm or less were used for RNA isolation.  
This is a directionally cloned, non-normalized library.  
Avg insert length: 1611; Primers: M13P and M13R;  
Antibiotic: 50 ug/ml Ampicillin; Primary Titer: 2.24E6 pfu  
total; Amplified Titer: 1.37E10 pfu/ml; Mass Excised  
Titer: 3.53E9 total; This library has been generated by  
the Floral Genome Project (FGP). We would like to thank  
David Lorence at the National Tropical Botanical Garden  
for providing plant material for library building. The  
Floral Genome Project is funded by NSF's Plant Genome  
Research Program (DBI-0115684). More information about the

project can be obtained at <http://fgp.bio.psu.edu>"

ORIGIN  
Query Match 100.0%; Score 15; DB 6; Length 685;  
Best Local Similarity 100.0%; Pred. No. 9.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 421 GACACCTATGTGCAC 435

RESULT 3  
AQ378906  
LOCUS  
DEFINITION  
AQ378906 698 bp DNA linear GSS 20-MAY-1999  
RPC111-16418.TV RPCI-11 Homo sapiens genomic clone RPCI-11-16418,  
genomic survey sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AQ378906.1 GI:4349929  
GSS.  
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and  
Venter,J.C.

REFERENCE  
AUTHORS  
1 (bases 1 to 698)  
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Venter,J.C.

TITLE  
Map Building  
Unpublished (1997)

JOURNAL  
COMMENT  
Other GSSs: RPC111-16418.TJ  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
([pieterdejong.med.buffalo.edu](mailto:pieterdejong.med.buffalo.edu)). Clones may be purchased from  
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from  
Research Genetics ([info@resgen.com](http://info@resgen.com)). BAC end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html)  
Seq primer: T7

Class: BAC ends

FEATURES  
source

Location/Qualifiers  
1..698  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7562791"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-16418"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_lib="RPCI-11"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPC111 Human Male BAC Library"

## ORIGIN

Query Match 100.0%; Score 15; DB 8; Length 698;  
Best Local Similarity 100.0%; Pred. No. 9.7e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCAC 15  
|||||  
Db 308 GACACCTATGTGCAC 322

RESULT 4  
BI086158  
LOCUS  
DEFINITION  
BI086158 711 bp mRNA linear EST 20-JUN-2001  
602870041T1 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5013011 3',

```

mRNA sequence.
ACCESSION BI086158 GI:14504488
VERSION BI086158.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 711)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/;
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1818 row: a column: 12
High quality sequence start: 9
High quality sequence stop: 163.
FEATURES
source
1..711
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5013011"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_98"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 100.0%; Score 15; DB 4; Length 711;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACACCTATGTGCAC 15
|||||
Db 411 GACACCTATGTGCAC 425

RESULT 5
AG480637 754 bp DNA linear GSS 04-JUN-2004
LOCUS AG480637
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-374P12.TJ, genomic survey
sequence.
ACCESSION AG480637
VERSION AG480637.1 GI:48187867
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE BAC end Sequences of Library MSMg01
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 754)
AUTHORS Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

mRNA sequence.
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@sc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@tc.riken.jp).
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@tc.riken.jp
PRIMERS
Sequencing: TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES
source
1..754
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-374P12.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match 100.0%; Score 15; DB 9; Length 754;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACACCTATGTGCAC 15
|||||
Db 165 GACACCTATGTGCAC 179

RESULT 6
BM322137 1030 bp mRNA linear EST 12-NOV-2002
LOCUS BM322137
DEFINITION RN79773 Sea urchin radial nerve Lambda Zap cDNA library
Helicoidaris erythrogramma cDNA 5', mRNA sequence.
ACCESSION BM322137
VERSION BM322137.1 GI:24901611
KEYWORDS EST.
SOURCE Helicoidaris erythrogramma
ORGANISM Helicoidaris erythrogramma
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida; Echinometridae;
Helicoidaris.
REFERENCE 1 (bases 1 to 1030)
AUTHORS Sly, B.J., Hazel, J.C., Popodi, E.M. and Raff, R.A.
TITLE Patterns of gene expression in the developing adult sea urchin
central nervous system reveal multiple domains and deep-seated
neural pentamery
JOURNAL Evol. Dev. 4 (3), 189-204 (2002)
MEDLINE 22049027
PUBMED 12054292
COMMENT Contact: R. A. Raff
Indiana Molecular Biology Institute and Department of Biology
Indiana University
1001 E. 3rd St., Bloomington, IN 47405, USA
Tel: 812 855 2791
Fax: 812 855 6705
Email: rraff@bio.indiana.edu
Seq primer: T3
High quality sequence stop: 1030
POLYA=No.
FEATURES
source
1..1030
/organism="Helicoidaris erythrogramma"
/mol_type="mRNA"
/db_xref="taxon:7634"

```

```

/tissue_type="Radial nerve"
/dev_stage="Adult"
/lab_host="E. coli"
/clone_lib="Sea urchin radial nerve Lambda Zap cDNA
library"
/notes="Vector: Bluescript SK; oligo dt priming from poly
A+ RNA, not directionally cloned"

ORIGIN
Query Match          100.0%; Score 15; DB 4; Length 1030;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTCAC 15
    |||||
Db 302 GACACCTATGTCAC 316

RESULT 7
BH908390/c
LOCUS
DEFINITION
SALK_047890.15.35.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_047890.15.35.x, genomic
survey sequence.
62 bp DNA linear GSS 04-SEP-2002
ACCESSION
BH908390
VERSION
BH908390.1 GI:22721323
KEYWORDS
GSS.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis;
1 (bases 1 to 62)
REFERENCE
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
FEATURES
Location/Qualifiers
1..62
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_047890.15.35.x"
/notes="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"

ORIGIN
Query Match          93.3%; Score 14; DB 8; Length 62;
Best Local Similarity 100.0%; Pred. No. 3e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14
    |||||
Db 58 GACACCTATGTGCA 45

/tissue_type="Radial nerve"
/dev_stage="Adult"
/lab_host="E. coli"
/clone_lib="Sea urchin radial nerve Lambda Zap cDNA
library"
/notes="Vector: Bluescript SK; oligo dt priming from poly
A+ RNA, not directionally cloned"

ORIGIN
Query Match          100.0%; Score 15; DB 4; Length 1030;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTCAC 15
    |||||
Db 302 GACACCTATGTCAC 316

RESULT 7
BH908390/c
LOCUS
DEFINITION
SALK_047890.15.35.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_047890.15.35.x, genomic
survey sequence.
62 bp DNA linear GSS 04-SEP-2002
ACCESSION
BH908390
VERSION
BH908390.1 GI:22721323
KEYWORDS
GSS.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis;
1 (bases 1 to 62)
REFERENCE
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.
FEATURES
Location/Qualifiers
1..62
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_047890.15.35.x"
/notes="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"

ORIGIN
Query Match          93.3%; Score 14; DB 8; Length 62;
Best Local Similarity 100.0%; Pred. No. 3e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14
    |||||
Db 58 GACACCTATGTGCA 45

```

```

RESULT 8
CK578709
LOCUS
DEFINITION
IST_W15_19848 AD-wrmcDNA library Caenorhabditis elegans cDNA 5',
similar to T05H10.6, mRNA sequence.
193 bp mRNA linear EST 16-JAN-2004
CK578709
CK578709.1 GI:40962377
EST.
Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 193)
Li,S., Armstrong,C.M., Bertin,N., Ge,H., Milstein,S., Boxem,M.,
Vidaian,P.O., Han,J.D., Chesneau,A., Hao,F., Goldberg,D.S., Li,N.,
Martinez,M., Rual,J.F., Lamesch,P., Xu,L., Tewari,M., Wong,S.L.,
Zhang,L.V., Beriz,G.F., Jacotot,L., Vaglio,P., Reboul,J.,
Hirozane-Kishikawa,T., Li,Q., Gabel,H.W., Elewa,A., Baumgartner,B.,
Rose,D.J., Yu,H., Bosak,S., Sequerra,R., Fraser,A., Baumgartner,B.,
Saxton,W.M., Stromer,S., Van Den Heuvel,S., Piano,P.,
Vandenhaute,J., Sardet,C., Gerstein,M., Doucette-Stamm,L.,
Gunsalus,K.C., Harper,J.W., Cusick,M.E., Roth,F.P., Hill,D.E. and
Vidal,M.
A Map of the Interactome Network of the Metazoan C. elegans
Science (2004) In press
Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
For the purpose of protein interaction mapping, we generated a C.
elegans cDNA library (AD-wrmcDNA) in which poly(dT)-primed reverse
transcribed cDNA are fused to the AD-encoding sequence of the yeast
transcription factor GAL4. cDNAs were generated and cloned into the
two hybrid vector pPC86 This Interacting Sequence Tag IST_W15_19848
(T05H10.6) interacts as a prey with the bait R13P6.9
PCR Primers
FORWARD: CGCGTTTGGATCCTACACGAGG
BACKWARD: GGAGACTTGACCAACCTCTGGCG
Insert Length: 193 Std Error: 157.00
Plate: 257 row: 09 column: C
Seq primer: CGCGTTTGGATCCTACACGAGG
High quality sequence stop: 192
POLYA=No.
FEATURES
Location/Qualifiers
1..193
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/dev_stage="embryos, L1, L2, L3, L4, adult, dauer"
/sex="male, hermaphrodite"
/clone_lib="AD-wrmcDNA library"
/notes="Vector: pPC86; For the purpose of protein
interaction mapping, we generated a C. elegans cDNA
library (AD-wrmcDNA) in which poly(dT)-primed reverse
transcribed cDNA are fused to the AD-encoding sequence of
the yeast transcription factor GAL4. This library was made
with poly(A)+ RNA isolated from mated populations of
wild-type (N2 strain) animals of all stages of development
including embryonic, larval (L1 to L4 stages) adults and
dauer. Approximately equal quantities of RNA from
different populations were acquired. cDNAs were generated
and cloned into the two hybrid vector pPC86. The library
contains ~3*10e7 clones. Reference - GATEWAY
recombinational cloning: application to the cloning of
large numbers of open reading frames or ORFeomes - Walhout
AJ, Temple GF, Brasch MA, Hartley JL, Lorson MA, van den
Heuvel S, Vidal M - Methods Enzymol. 2000;328:575-92"
ORIGIN

```

Query Match 93.3%; Score 14; DB 7; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+03;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15  
 |||||  
 Db 45 ACACCTATGTGCAC 58

RESULT 9  
 CD340276 255 bp mRNA linear EST 17-SEP-2003  
 LOCUS CD340276  
 DEFINITION MPMP336 Strongylocentrotus purpuratus cDNA clone  
 CALTP536P0747;MP1\_536\_47P7 5', mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS

SOURCE  
 ORGANISM  
 Strongylocentrotus purpuratus

Strongylocentrotus purpuratus  
 Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 Echinoidea; Euechinoidea; Echinacea; Echinoidea;  
 Strongylocentrotidae; Strongylocentrotus.

REFERENCE  
 1 (bases 1 to 255)  
 Poustka, A.J., Groth, D., Hennig, S., Thamm, S., Cameron, A., Beck, A.,  
 Reinhardt, R., Herwig, R., Panopoulou, G. and Lehrach, H.  
 Generation, annotation, evolutionary analysis, and database  
 integration of 20,000 unique sea urchin EST clusters  
 Genome Res. 13 (12), 2736-2746 (2003)

JOURNAL  
 COMMENT  
 Contact: Poustka AJ  
 Laboratory 145, Dept. Lehrach  
 Max-Planck-Institut fuer Molekulare Genetik  
 Ihnestr.63-73, D-14195 Berlin, Germany  
 Tel: +49 30 8413 1235  
 Fax: +49 30 8413 1128  
 Email: poustka@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting  
 (ONF) to reduce sequencing redundancy. According to the ONF  
 procedure, clones that display the same hybridisation matrix with a  
 battery of 200 8mer oligonucleotides are grouped into clusters. One  
 clone per ONF cluster is selected for sequencing. The size of each  
 cluster is an indicator of the frequency of a transcript in the  
 analysed library. The cluster size as well as the coordinates of  
 the other clones assigned to the same ONF cluster as the clone from  
 which the above EST is generated is available at the sea urchin  
 project web site at: [http://www.molgen.mpg.de/ag\\_seaurchin/](http://www.molgen.mpg.de/ag_seaurchin/). cDNA  
 clones and filters are distributed via the Resource Center/Primary  
 Database of the German Human Genome Project (<http://www.rzpd.de/>)  
 PCR Primers  
 FORWARD: 5' CCCGAGGCTTTACACTTTATGCTTCCGGCTCG 3' (M13RSP) 5'-seq  
 BACKWARD: 5' GCTATTACGCAGCTGGCGAAAGGGGATGTG 3' (M13FSP) 3'-seq  
 Seq primer: 5'-CCGGTCCGGAATTCGGGT-3' pSport3/86  
 High quality sequence stop: 255.

FEATURES  
 source

Location/Qualifiers  
 1..255  
 /organism="Strongylocentrotus purpuratus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7668"  
 /clone="CALTP536P0747;MP1\_536\_47P7"  
 /tissue\_type="whole embryo"  
 /dev\_stage="embryonic 40hr"  
 /lab\_host="E.coli, XL1 blue"  
 /clone\_lib="Sea urchin embryo 40hr gastrula stage cDNA  
 library MPMP336"  
 /note="Vector: pSport1; Site 1: NotI; Site 2: SalI; Random  
 primed and directionally cloned in pSport1 vector using a  
 NotI (5'-pGATGTTCTAGATCGGAGCGCGGCC (T)15-3' and a  
 SalI 5'-TCGACCCAGCGCTCCG-3' adapters (Gibco BRL)"

Query Match 93.3%; Score 14; DB 6; Length 255;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+03;

ORIGIN

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 ACACCTATGTGCAC 15  
 |||||  
 Db 80 ACACCTATGTGCAC 93

RESULT 10  
 AA558239/c

LOCUS AA558239  
 DEFINITION n134g05.s1 NCI CGAP HSC1 Homo sapiens cDNA clone IMAGE:1042616  
 similar to TR:G409466 G409466 CGI PROTEIN PRECURSOR. [2] TR:G296164  
 ; mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS

SOURCE  
 ORGANISM  
 Homo sapiens (human)

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 1 (bases 1 to 273)

AUTHORS  
 TITLE  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL  
 COMMENT  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapsb@mail.nih.gov

Tissue Procurement: Donald Orlic, Ph.D.  
 cDNA Library Preparation: David B. Krizman, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert Length: 393 Std Error: 0.00  
 Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES  
 source

Location/Qualifiers  
 1..273  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1042616"  
 /tissue\_type="stem cells"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP HSC1"  
 /note="Organ: bone marrow; Vector: pAMP10; mRNA made from  
 flow-sorted CD34+/CD38- hematopoietic stem cells, cDNA  
 made by oligo-dT priming. Non-directionally cloned.  
 Size-selected on agarose gel, average insert size 600 bp.  
 cDNA Library Preparation: David B. Krizman, Ph.D.  
 Reference: Krizman et al. (1996) Cancer Research  
 56:5380-5383. cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing  
 Center"

ORIGIN

Query Match 93.3%; Score 14; DB 1; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+03;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14  
 |||||  
 Db 90 GACACCTATGTGCA 77

RESULT 11  
 AI702394

LOCUS AI702394  
 DEFINITION t267b08.x1 NCI CGAP Panl Homo sapiens cDNA clone IMAGE:2293623 3'  
 similar to contains element MER22 repetitive element ;, mRNA  
 sequence.

ACCESSION  
 VERSION  
 AI702394.1 GI:4990294







Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae; Cichlidae; Haplochromis.

1 (bases 1 to 353)  
Watanabe,M., Kobayashi,N., Shin-i,T., Kohara,Y. and Okada,N.  
Orf sequences of cichlid in Lake Victoria are essentially same  
Unpublished (2004)  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES  
source  
1..353  
Location/Qualifiers  
/organism="Haplochromis sp. 'red tail sheller'"  
/mol\_type="mRNA"  
/db\_xref="taxon:257976"  
/clone="no574b02"  
/tissue\_type="jaw"  
/dev\_stage="varied"  
/clone\_lib="HRESt library"

## ORIGIN

Query Match 93.3%; Score 14; DB 4; Length 353;  
Best Local Similarity 100.0%; Pred. No. 3.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15

|||||

Db 93 ACACCTATGTGCAC 80

## RESULT 17

CK926149/c  
LOCUS  
CK926149.1 354 bp mRNA linear EST 22-APR-2004  
DEFINITION  
p3fmgc\_017775 Normalized Magnaporthe grisea cDNA PRS423 library  
Magnaporthe grisea cDNA clone p3fmgc\_017775, mRNA sequence.

CK926149

EST.

CK926149.1 GI:45394501

SOURCE  
Magnaporthe grisea (anamorph: Pyricularia grisea)

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.  
1 (bases 1 to 354)

Dong,H., Peng,X., Chen,B., Li,Y. and Li,D.  
Large-scale identification of ESTs from Magnaporthe grisea by  
normalized cDNA library sequencing (2004b)

Unpublished (2004)  
JOURNAL  
COMMENT  
Contact: Debao Li  
Bioinformatics and Gene Network Research Group  
Zhejiang University, China Agricultural University, Guangxi  
University  
Kaixuan Road 268#, Hangzhou, Zhejiang, 310029, china  
Tel: 0086-571-86961525  
Fax: 0086-571-86971183  
Email: webmaster@estarray.org, URL: http://www.estarray.org

## FEATURES

source  
1..354  
Location/Qualifiers  
/organism="Magnaporthe grisea"  
/mol\_type="mRNA"  
/db\_xref="taxon:149305"  
/clone="p3fmgc\_017775"  
/tissue\_type="Mycellium, conidium, germinating conidium,  
swelling appressorium, mature appressorium, penetration  
peg"  
/dev\_stage="Mycellium, conidium, germinating conidium,  
swelling appressorium, mature appressorium, penetration  
peg"

## ORIGIN

Query Match 93.3%; Score 14; DB 7; Length 354;  
Best Local Similarity 100.0%; Pred. No. 3.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15

|||||

Db 253 ACACCTATGTGCAC 240

## RESULT 18

BB841157  
LOCUS  
BB841157 360 bp mRNA linear EST 21-NOV-2001  
DEFINITION  
BB841157 RIKEN full-length enriched, 6 days neonate spleen Mus  
musculus cDNA clone F42005J19 5', mRNA sequence.

BB841157

BB841157.1 GI:17041888

EST.

SOURCE  
Mus musculus (house mouse)

ORGANISM  
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 360)

Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,  
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,  
Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Konda,M.,  
Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,  
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,  
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
Tanaka,T., Tomaru,A., Toyai,T., Wataniki,A., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.  
2001)

Unpublished (2001)

CONTACT: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,  
Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.  
and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,  
Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (http://genome.gsc.riken.go.jp) for  
further details.

e mouse tissues.

Location/Qualifiers  
1..360  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"

```

/clone="F420005J19"
/tissue type="spleen"
/dev_stage="6 days neonate"
/clone_lib="RIKEN full-length enriched, 6 days neonate
spleen"

ORIGIN
Query Match          93.3%; Score 14; DB 2; Length 360;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACACCTATGTGCA 14
|||||
103 GACACCTATGTGCA 116

RESULT 19
CL327998          375 bp   DNA   linear   GSS 19-AUG-2004
LOCUS             CL327998
DEFINITION        CL327998.1 r RPCI-44 Sus scrofa genomic clone RPCI44_246D14,
genomic survey sequence.
ACCESSION         CL327998
VERSION           CL327998.1 GI:51379968
KEYWORDS          GSS.
SOURCE            Sus scrofa (pig)
ORGANISM          Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE         1 (bases 1 to 375)
AUTHORS           Rogatcheva,M.B., Meyers,S., He,W., Larkin,D.M., Marron,B.M.,
Beever,J.E. and Schook,L.B.
TITLE             Piggy-BACing the Human Genome: Constructing a Porcine Physical Map
Through Comparative Genomics
JOURNAL            Unpublished (2004)
COMMENT           Other GSSs: RPCI44_246D14.f
Contact: Lawrence B. Schook
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 265 5326
Fax: 217 244 5617
Email: schook@uiuc.edu
Clones are derived from the porcine BAC library RPCI-44
(http://www.bacpac.chori.org/porcine242.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@chori.org).
Clones may be purchased from BACPAC Resources
(http://BACPACresources.chori.org). This work was undertaken as part
of the International Swine Genome Sequencing Consortium by
University of Illinois at Urbana Champaign, USA with funds provided
by grant No. AG2002-34480-11828 from USDA-CSREES and
AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing
Initiative)
Plate: 246 row: D column: 14
Seq primer: SP6
Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..375
                     /organism="Sus scrofa"
                     /mol_type="genomic DNA"
                     /strain="four pigs (breed: 37.5% Yorks Landrace and 25%
                     Meishan)"
                     /db_xref="taxon:9823"
                     /clones="RPCI44_246D14"
                     /sex="male"
                     /cell_type="blood"
                     /clone_lib="RPCI-44"
                     /note="Vector: pTARBAC2; Site 1: EcoRI; Site 2: EcoRI;
                     porcine male BAC library produced by Pieter de Jong"

ORIGIN
Query Match          93.3%; Score 14; DB 9; Length 375;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACACCTATGTGCA 14
|||||
103 GACACCTATGTGCA 116

RESULT 20
BB817276          381 bp   mRNA   linear   EST 19-NOV-2001
LOCUS             BB817276
DEFINITION        BB817276 RIKEN full-length enriched, lung RCB-0558 LLC cDNA Mus
musculus cDNA clone G730038003 3', mRNA sequence.
ACCESSION         BB817276
VERSION           BB817276.1 GI:16989905
KEYWORDS          EST.
SOURCE            Mus musculus (house mouse)
ORGANISM          Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE         1 (bases 1 to 381)
AUTHORS           Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
Ihii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
Okazaki,Y., Okido,T., Saito,K., Sakai,C., Sakai,K., Sakazume,N.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Tanaka,T., Tomaru,A., Toyota,T., Watahiki,A., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE             RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
JOURNAL            Unpublished (2001)
COMMENT           Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gscl.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
FEATURES             Location/Qualifiers
     source           1..381
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clones="G730038003"
                     /tissue_type="lung"
                     /call_line="RCB-0558 LLC"
                     /clone_lib="RIKEN full-length enriched, lung RCB-0558 LLC
                     cDNA"

ORIGIN
Query Match          93.3%; Score 14; DB 2; Length 381;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;

```

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14

Db 13 GACACCTATGTGCA 26

RESULT 21

BY670540

LOCUS

DEFINITION

BY670540 RIKEN full-length enriched, 14.5 days embryo df/df  
Rathke's pouches Mus musculus cDNA clone K820013C23 3', mRNA

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 GACACCTATGTGCA 14  
|||||  
13 GACACCTATGTGCA 26  
BY670540 383 bp mRNA linear EST 16-DEC-2002  
Rathke's pouches Mus musculus cDNA clone K820013C23 3', mRNA  
sequence.  
BY670540  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 383)  
Nikaido, I., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,  
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,  
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,  
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,  
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K.,  
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,  
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,  
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,  
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,  
Rogers, J., Birney, E. and Hayashizaki, Y.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
12466851  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,  
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,  
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,  
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,  
Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and  
Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues

Tissues were provided by Michelle Brinkmeier and Sally Camper (Dept.  
Human Genetics University of Michigan Medical School 4301  
MSRB 3 1500 W. Medical Center Dr. Ann Arbor, MI 48109-0638 USA )  
whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

FEATURES  
source

Location/Qualifiers  
1. 383  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="K820013C23"  
/tissue\_type="Rathke's pouches"  
/dev\_stage="14.5 days embryo df/df"  
/clone\_lib="RIKEN full-length enriched, 14.5 days embryo  
df/df Rathke's pouches"

ORIGIN

Query Match 93.3%; Score 14; DB 6; Length 383;  
Best Local Similarity 100.0%; Pred. No. 3.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14

Db 34 GACACCTATGTGCA 47

RESULT 22

BY615493/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BY615493 391 bp mRNA linear EST 15-DEC-2002  
BY615493 RIKEN full-length enriched, visual cortex Mus musculus  
cDNA clone K330010M09 3', mRNA sequence.  
BY615493  
GI:26950675  
Mus musculus (house mouse)  
Mus musculus  
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 391)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,  
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,  
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,  
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,  
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K.,  
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,  
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,  
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

Arakawa, T., Fukuda, S., Hara, A., Haehizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

**TITLE** Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

**JOURNAL** Nature 420, 563-573 (2002)

**MEDLINE** 22354683

**PUBMED** 12466851

**COMMENT** Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ono, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Michela Fagioli and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hiroseawa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

# FEATURES

source  
1. .391  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="K330010M09"  
/tissue\_type="visual cortex"  
/clone\_lib="RIKEN full-length enriched, visual cortex"

# ORIGIN

Query Match 93.3%; Score 14; DB 6; Length 391;  
Best Local Similarity 100.0%; Pred. No. 3.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14

|||||

Db 36 GACACCTATGTGCA 23

# RESULT 23

AQ597571

**LOCUS**

**DEFINITION** HS 2090 A2 H05 T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2090 Col=10 Row=O, genomic survey sequence.

**ACCESSION** AQ597571

**VERSION** AQ597571.1 GI:5028783

# KEYWORDS

**SOURCE**

**ORGANISM**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**MEDLINE**

**PUBMED**

**COMMENT**

**CONTACT**

**EMAIL**

**URL**

**ABSTRACT**

**KEYWORDS**

**ORGANISM**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**COMMENT**

**CONTACT**

**EMAIL**

**URL**

**ABSTRACT**

**KEYWORDS**

**ORGANISM**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**COMMENT**

**CONTACT**

**EMAIL**

**URL**

**ABSTRACT**

**KEYWORDS**

**ORGANISM**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**COMMENT**

**CONTACT**

**EMAIL**

**URL**

**ABSTRACT**

**KEYWORDS**

**ORGANISM**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**COMMENT**

**CONTACT**

**EMAIL**

**URL**

**ABSTRACT**

**KEYWORDS**

**ORGANISM**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**COMMENT**

**CONTACT**

**EMAIL**

**URL**

**ABSTRACT**

**KEYWORDS**

**ORGANISM**

**REFERENCE**

**AUTHORS**

**TITLE**

**JOURNAL**

**COMMENT**

**CONTACT**

**EMAIL**

**URL**

**ABSTRACT**

**KEYWORDS**

**ORGANISM**

# GSS.

Homo sapiens (human)

Homo sapiens

Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 397)

Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

9380589

10449764

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: <http://www.htsc.washington.edu>

Plate: 2090 row: O column: 10

Seq primer: T7

Class: BAC ends

High quality sequence stop: 397.

Location/Qualifiers

1. .397

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone="Plate=2090 Col=10 Row=O"

/sex="male"

/clone\_lib="CIT Approved Human Genomic Sperm Library D"

/notes="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

Query Match 93.3%; Score 14; DB 8; Length 397;

Best Local Similarity 100.0%; Pred. No. 3.5e+03;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15

|||||

Db 151 ACACCTATGTGCAC 164

# RESULT 24

AA632057

**LOCUS**

**DEFINITION** np70601.s1 NCI\_CGAP\_Br2 Homo sapiens cDNA clone IMAGE:1131672 3', mRNA sequence.

AA632057

**ACCESSION**

AA632057.1 GI:2555471

**KEYWORDS**

**SOURCE**

**ORGANISM**

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 399)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgap@ncic.nih.gov](mailto:cgap@ncic.nih.gov)

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 1509 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 388.

## FEATURES

source

```
1..399
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1131672"
/sex="female, pooled"
/tissue_type="breast"
/lab_host="DH10B"
/clone_lib="NCI CGAP Br2"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from pooled bulk
breast tumor tissue, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. This library is the normalized version of
NCI CGAP Br1.1. Library was constructed by Bento Soares
and M. Fátima Bonaldo."
```

## ORIGIN

```
Query Match      93.3%; Score 14; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 GACACCTATGTGCA 14

Db 39 GACACCTATGTGCA 52

RESULT 25  
AW354829/c

```
LOCUS      37769 MARC 2P1G Sus scrofa cDNA 5', mRNA linear EST 09-JUL-2000
DEFINITION      37769 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION      AW354829
VERSION      AW354829.1 GI:6853819
```

KEYWORDS EST.

SOURCE Sus scrofa (pig)

ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

1 (bases 1 to 402)

Fahrenkrug S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,

Vallet, J., Wise, T., Rohrer, G.A., Perlea, G., Sultana, R.,

Quackenbush, J., and Keele, J.W.

Porcine gene discovery by normalized cDNA-library sequencing and

EST cluster assembly

Mamm. Genome 13 (8), 475-478 (2002)

22213789

12226715

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smitht@mail.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred

v0.980904.e. Vector identified by cross\_match with the -minscore 20

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACACGTATGACCAT

BACKWARD: GTTTCCCGAGTCACGACG

Plate: 16 row: M column: 11

Seq primer: ATTAGGTGACACTATAG.

## FEATURES

source

1..402

/organism="Sus scrofa"

/mol\_type="mRNA"

```
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2P1G"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: Salt;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
```

## ORIGIN

```
Query Match      93.3%; Score 14; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 2 ACACCTATGTGCAC 15

Db 46 ACACCTATGTGCAC 33

RESULT 26

BY624835/c

LOCUS

```
DEFINITION      BY624835 403 bp mRNA linear EST 15-DEC-2002
cDNA clone K33033E23 3', mRNA sequence.
```

ACCESSION

BY624835

EST.

KEYWORDS Mus musculus (house mouse)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 403)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schönbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,

Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,

Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,

Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,

Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W.J., Perlea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,

Sandelin, A., Schneider, C., Sempie, C.A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,

Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,

Rogers, J., Birney, E., and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,

Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,

Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Wataniki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Michela Fagiolini and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hirotsawa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES source  
1. .403  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="K33033E23"  
/tissue\_type="visual cortex"  
/clone\_lib="RIKEN full-length enriched, visual cortex"

ORIGIN

Query Match 93.3%; Score 14; DB 6; Length 403;  
Best Local Similarity 100.0%; Pred. No. 3.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14  
|||||  
Db 47 GACACCTATGTGCA 34

RESULT 27  
AQ204278  
LOCUS  
DEFINITION  
HS\_3113\_B2\_A05\_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3113 Col=10 Row=B, genomic survey sequence.  
ACCESSION  
AQ204278.1 GI:3614848  
VERSION  
AQ204278.1  
KEYWORDS  
GSS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 404)  
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and Hood, L.  
TITLE  
Sequencing-tagged connectors: A sequence approach to mapping and scanning the human genome  
JOURNAL  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
MEDLINE  
93380589  
PUBMED  
10449764  
COMMENT  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618

FEATURES source  
1. .404  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="plate=3113 Col=10 Row=B"  
/sex="male"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

## ORIGIN

Query Match 93.3%; Score 14; DB 8; Length 404;  
Best Local Similarity 100.0%; Pred. No. 3.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACCTATGTGCAC 15  
|||||  
Db 220 ACACCTATGTGCAC 233

## RESULT 28

H25220/c  
LOCUS  
DEFINITION  
H25220 r1 Soares infant brain IN1B Homo sapiens cDNA clone IMAGE:52526 5', mRNA sequence.  
ACCESSION  
H25220.1 GI:894343  
VERSION  
H25220.1  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 406)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.  
TITLE  
The WashU-Merck EST Project  
JOURNAL  
Unpublished (1995)  
COMMENT  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
Insert Size: 2103

High quality sequence stops: 290 Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert Length: 2103 Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 290.

FEATURES source  
1. .406  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:425462"  
/db\_xref="taxon:9606"  
/clone="IMAGE:52526"  
/sex="female"  
/dev\_stage="73 days post natal"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares infant brain IN1B"  
/note="Organ: whole brain; Vector: Lafmid BA; Site 1: Not I; Site 2: Hind III; 1st strand cDNA was primed with a Not

I - oligo(dT) primer [5'  
AATGGGAATTCGGCCGCGAGGAATTTT 3'];  
double-stranded cDNA was ligated to Hind III adaptors  
(Pharmacia), digested with Not I and directionally cloned  
into the Not I and Hind III sites of the pLafmid BA vector.  
Library went through one round of normalization. Library  
constructed by Bento Soares and M.Fatima Bonaldo."

## ORIGIN

Query Match 93.3%; Score 14; DB 7; Length 406;  
Best Local Similarity 100.0%; Pred. No. 3.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14  
|||||  
Db 101 GACACCTATGTGCA 88

## RESULT 29

CB770156/c  
LOCUS  
DEFINITION CB770156 409 bp mRNA linear EST 16-MAY-2003  
srpb2-00130-h3-A srpb2 (10220) Rattus norvegicus cDNA clone  
CB770156  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

## REFERENCE

1 (bases 1 to 409)  
Angen Rat EST Program.  
TITLE  
JOURNAL  
COMMENT  
Contact: Dan Fitzpatrick  
Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00130 row: h column: 3.

## FEATURES

source

1..409  
Location/Qualifiers  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="srpb2-00130-h3"  
/tissue\_type="prostate tissue"  
/clone\_lib="srpb2 (10220)"  
/notes="Vector: pSPORI1; Site\_1: SalI; Site\_2: NotI; rat  
prostate normalized double selected poly(A+) mRNA size  
fraction > 1 kb"

## ORIGIN

Query Match 93.3%; Score 14; DB 6; Length 409;  
Best Local Similarity 100.0%; Pred. No. 3.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15  
|||||  
Db 263 ACACCTATGTGCAC 250

## RESULT 30

BQ296291  
LOCUS  
DEFINITION BQ296291 420 bp mRNA linear EST 05-JUL-2004  
san89f05.v2 Gm-c1052 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-c1052-6802 5' similar to TR:049929 O49929 POSE PROTEIN OF 24 KD  
/, mRNA sequence.

ACCESSION BQ296291  
VERSION BQ296291.1 GI:20811813  
KEYWORDS  
SOURCE  
ORGANISM

Glycine max (soybean)

REFERENCE  
AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 420)  
Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,  
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,I., Cardenas,M.,  
McCann,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
Unpublished (1999)

TITLE  
JOURNAL  
COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this  
clone is listed in the 'Other ESTs on clone' field. This clone is  
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD  
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)  
Seq primer: -4ORP from Gibco  
High quality sequence stop: 134.

FEATURES  
source

1..420  
Location/Qualifiers  
/organism="Glycine max"  
/mol\_type="mRNA"  
/cultivar="Harosoy"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-c1052-6802"  
/tissue\_type="whole seedlings of greenhouse grown plants"  
/dev\_stage="1 week old"  
/lab\_host="DH10B"  
/clone\_lib="Gm-c1052"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The Harosoy NIL was constructed and seed was  
provided by Dr. J. Specht, University of Nebraska  
(Shoemaker and Specht, 1995). The cDNA library was  
constructed from mRNA isolated from whole seedlings of 1  
week old greenhouse grown plants. Complementary DNA was  
synthesized from mRNA using a primer consisting of a  
poly(dT) sequence with a XhoI restriction site and a 3'  
anchor. EcoRI adapters were ligated to the blunt-ended  
cDNA fragments followed by XhoI digestion. The cDNA  
fragments were directionally cloned into the EcoRI-XhoI  
restriction site of the pBluescript vector. The ligated  
cDNA fragments were transformed into DH10B host cells  
(GibcoBRL). The library was constructed in cooperation  
with Dr. Paul Keim's laboratory at Northern Arizona  
University."

## ORIGIN

Query Match 93.3%; Score 14; DB 5; Length 420;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACCTATGTGCAC 15  
|||||  
Db 371 ACACCTATGTGCAC 384

## RESULT 31

BY370371/c  
LOCUS  
DEFINITION BY370371 420 bp mRNA linear EST 09-JUL-2003  
musculus cDNA clone G43008N03 3', mRNA sequence.  
ACCESSION BY370371  
VERSION BY370371.1 GI:26599859  
KEYWORDS  
SOURCE  
Mus musculus (house mouse)



```

ORGANISM      Mus musculus
REFERENCE
AUTHORS       Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
              1 (bases 1 to 420)
              Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
              Kikado, I., Otsu, N., Saito, R., Suzuki, H., Yananaka, I.,
              Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
              Schonbach, C., Gojohori, T., Baldarelli, R., Hill, D.P., Bult, C.,
              Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
              Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
              Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
              Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
              Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
              Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
              Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
              Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
              Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
              Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G.,
              Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
              Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
              Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
              Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
              Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
              Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I.,
              Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
              Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
              Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
              Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
              Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
              Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
              Rogers, J., Birney, E. and Hayashizaki, Y.
              Analysis of the mouse transcriptome based on functional annotation
              of 60,770 full-length cDNAs
              Nature 420, 563-573 (2002)
              22354683
              12466851
              Contact: Yoshihide Hayashizaki
              Laboratory for Genome Exploration Research Group, RIKEN Genomic
              Sciences Center (GSC), Yokohama Institute
              The Institute of Physical and Chemical Research (RIKEN)
              1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
              Tel: 81-45-503-9222
              Fax: 81-45-503-9216
              Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
              Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
              Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
              Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
              Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
              Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and
              Hayashizaki, Y. Direct Submission
              Computational Analysis of Full-length Mouse cDNAs Compared with
              Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
              Normalization and subtraction of cap-trapper-selected cDNAs to
              prepare full-length cDNA libraries for rapid discovery of new
              genes. Genome Res. 10 (10), 1617-1630 (2000)
              RIKEN integrated sequence analysis (RISA) system--384-format
              sequencing pipeline with 384 multicapillary sequencer. Genome Res.
              10 (11), 1757-1771 (2000)
              Computer-based methods for the mouse full-length cDNA
              encyclopedia: real-time sequence clustering for construction of a
              nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
              cDNA library was prepared and sequenced in Mouse Genome
              Encyclopedia Project of Genome Exploration Research Group in Riken
              Genomic Sciences Center and Genome Science Laboratory in RIKEN.
              Division of Experimental Animal Research in Riken contributed to
              prepare mouse tissues.
              Please visit our web site (http://genome.gsc.riken.go.jp) for
              further details.
              Location/Qualifiers
              1. .420
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="BALB/c"

FEATURES
source
  1. .420
  /organism="Mus taurus"
  /mol_type="genomic DNA"
  /strain="breed: Hereford"
  /db_xref="taxon:9913"
  /clone="CH240_155L09"
  /sex="Male"
  /cell_type="Blood"
  /clone_lib="CHORI-240"
  /notes="Vector: PTAREBAC1.3; Site 1: MboI; Site 2: MboI;
  Hereford bull LI Domino 99375; CHORI-240 Bovine BAC

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
  Contact: Caetano AR
  Other_GSSs: CH240_155L09.TV
  Department of Biotechnology
  Embrapa Recursos Geneticos e Biotecnologia
  Parque Estacao Biologica, Final Av. W/5 Norte, Brasilia-DF C.P.
  02372, 70770-900 Brasil
  Tel: 55 61 448 4778
  Fax: 55 61 340 3658
  Email: caetano@cenargen.embrapa.br
  Clones are derived from the bovine BAC library CHORI-240
  (http://www.chori.org/bacpac/bovine240.htm).
  Bases shown have phred quality value equal to or higher than 20.
  Bases with quality value below 20 were masked with 'N'.
  For BAC library availability, please contact Pieter de Jong
  (pdejong@mail.choi.org).
  Clones may be purchased from BACPAC Resources
  (http://www.chori.org/bacpac/ordering information.htm).
  This work was undertaken as part of the International Bovine BAC
  Mapping Consortium (IBBMC) by Embrapa Recursos Geneticos e
  Biotecnologia with financing from Conselho Nacional de
  Desenvolvimento Cientifico e Tecnol6gico (CNPq), Brazil
  Plate: 155 row: L column: 09
  Seq primer: SP6
  Class: BAC ends
  High quality sequence stop: 420.
  Location/Qualifiers
  1. .420
  /organism="Bos taurus"
  /mol_type="genomic DNA"
  /strain="breed: Hereford"
  /db_xref="taxon:9913"
  /clone="CH240_155L09"
  /sex="Male"
  /cell_type="Blood"
  /clone_lib="CHORI-240"
  /notes="Vector: PTAREBAC1.3; Site 1: MboI; Site 2: MboI;
  Hereford bull LI Domino 99375; CHORI-240 Bovine BAC

ORIGIN
Query Match 93.3%; Score 14; DB 5; Length 420;
Best Local Similarity 100.0%; Pred. No. 3.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14
  |||||
Db 87 GACACCTATGTGCA 74

RESULT 32
CG985509/c
LOCUS
DEFINITION
  CG985509 420 bp DNA linear GSS 15-DEC-2003
  CH240_155L09.TJ CHORI-240 Bos taurus genomic clone CH240_155L09,
  genomic survey sequence.
ACCESSION
  CG985509
VERSION
  CG985509.1 GI:39911288
KEYWORDS
  GSS.
SOURCE
  Bos taurus (cow)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
  Bovinae; Bos.
  1 (bases 1 to 420)
  Costa, J.N., Mota, M. and Caetano, A.R.
  Brazil's Contribution to End-Sequencing the Bovine BAC Library
  CHORI-240
  Unpublished (2003)
  Other_GSSs: CH240_155L09.TV
  Contact: Caetano AR
  Department of Biotechnology
  Embrapa Recursos Geneticos e Biotecnologia
  Parque Estacao Biologica, Final Av. W/5 Norte, Brasilia-DF C.P.
  02372, 70770-900 Brasil
  Tel: 55 61 448 4778
  Fax: 55 61 340 3658
  Email: caetano@cenargen.embrapa.br
  Clones are derived from the bovine BAC library CHORI-240
  (http://www.chori.org/bacpac/bovine240.htm).
  Bases shown have phred quality value equal to or higher than 20.
  Bases with quality value below 20 were masked with 'N'.
  For BAC library availability, please contact Pieter de Jong
  (pdejong@mail.choi.org).
  Clones may be purchased from BACPAC Resources
  (http://www.chori.org/bacpac/ordering information.htm).
  This work was undertaken as part of the International Bovine BAC
  Mapping Consortium (IBBMC) by Embrapa Recursos Geneticos e
  Biotecnologia with financing from Conselho Nacional de
  Desenvolvimento Cientifico e Tecnol6gico (CNPq), Brazil
  Plate: 155 row: L column: 09
  Seq primer: SP6
  Class: BAC ends
  High quality sequence stop: 420.
  Location/Qualifiers
  1. .420
  /organism="Bos taurus"
  /mol_type="genomic DNA"
  /strain="breed: Hereford"
  /db_xref="taxon:9913"
  /clone="CH240_155L09"
  /sex="Male"
  /cell_type="Blood"
  /clone_lib="CHORI-240"
  /notes="Vector: PTAREBAC1.3; Site 1: MboI; Site 2: MboI;
  Hereford bull LI Domino 99375; CHORI-240 Bovine BAC

```

library (Male) produced by Pieter de Jong"

ORIGIN

Query Match 93.3%; Score 14; DB 9; Length 420;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14  
|||||  
Db 196 GACACCTATGTGCA 183

RESULT 33  
N58741  
LOCUS  
DEFINITION yv74d12.g1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
IMAGE:248471 3', mRNA sequence.

ACCESSION N58741  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 421)  
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,  
Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,  
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,  
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,  
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,  
Trevisan,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.,  
and Marra,M.

TITLE Generation and analysis of 280,000 human expressed sequence tags  
JOURNAL Genome Res. 6 (9), 807-828 (1996)  
MEDLINE 97044478  
PUBMED 8889549

CONTACT: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 966 Std Error: 0.00  
Seq primer: mob.REGA+ET

High quality sequence stop: 1.  
Location/Qualifiers  
1. .421  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3797717"  
/db\_xref="taxon:9606"  
/clone="IMAGE:248471"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/note="Organ: Liver and Spleen; Vector: pRT3D (Pharmacia)  
with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5', AACTGGAAGATTAATTAAGATCTTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pRT3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Patima Bonaldo."

FEATURES  
source  
1. .421  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3797717"  
/db\_xref="taxon:9606"  
/clone="IMAGE:248471"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/note="Organ: Liver and Spleen; Vector: pRT3D (Pharmacia)  
with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5', AACTGGAAGATTAATTAAGATCTTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pRT3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Patima Bonaldo."

ORIGIN

Query Match 93.3%; Score 14; DB 7; Length 421;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14  
|||||  
Db 91 GACACCTATGTGCA 78

RESULT 35  
CR555382/c  
LOCUS  
DEFINITION DKF2p459D2446 r1 459 (synonym: pcor1Pongo pygmaeus cDNA clone

library (Male) produced by Pieter de Jong"

ORIGIN

Query Match 93.3%; Score 14; DB 9; Length 420;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14  
|||||  
Db 196 GACACCTATGTGCA 183

RESULT 33  
N58741  
LOCUS  
DEFINITION yv74d12.g1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
IMAGE:248471 3', mRNA sequence.

ACCESSION N58741  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 421)  
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,  
Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,  
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,  
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,  
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,  
Trevisan,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.,  
and Marra,M.

TITLE Generation and analysis of 280,000 human expressed sequence tags  
JOURNAL Genome Res. 6 (9), 807-828 (1996)  
MEDLINE 97044478  
PUBMED 8889549

CONTACT: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 966 Std Error: 0.00  
Seq primer: mob.REGA+ET

High quality sequence stop: 1.  
Location/Qualifiers  
1. .421  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3797717"  
/db\_xref="taxon:9606"  
/clone="IMAGE:248471"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/note="Organ: Liver and Spleen; Vector: pRT3D (Pharmacia)  
with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5', AACTGGAAGATTAATTAAGATCTTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pRT3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Patima Bonaldo."

FEATURES  
source  
1. .421  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3797717"  
/db\_xref="taxon:9606"  
/clone="IMAGE:248471"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/note="Organ: Liver and Spleen; Vector: pRT3D (Pharmacia)  
with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5', AACTGGAAGATTAATTAAGATCTTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pRT3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Patima Bonaldo."

ORIGIN

Query Match 93.3%; Score 14; DB 7; Length 421;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14  
|||||  
Db 91 GACACCTATGTGCA 78

RESULT 35  
CR555382/c  
LOCUS  
DEFINITION DKF2p459D2446 r1 459 (synonym: pcor1Pongo pygmaeus cDNA clone

library (Male) produced by Pieter de Jong"

ORIGIN

Query Match 93.3%; Score 14; DB 9; Length 420;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACCTATGTGCA 14  
|||||  
Db 196 GACACCTATGTGCA 183

RESULT 33  
N58741  
LOCUS  
DEFINITION yv74d12.g1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
IMAGE:248471 3', mRNA sequence.

ACCESSION N58741  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 421)  
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,  
Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,  
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,  
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,  
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,  
Trevisan,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.,  
and Marra,M.

TITLE Generation and analysis of 280,000 human expressed sequence tags  
JOURNAL Genome Res. 6 (9), 807-828 (1996)  
MEDLINE 97044478  
PUBMED 8889549

CONTACT: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 966 Std Error: 0.00  
Seq primer: mob.REGA+ET

High quality sequence stop: 1.  
Location/Qualifiers  
1. .421  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3797717"  
/db\_xref="taxon:9606"  
/clone="IMAGE:248471"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/note="Organ: Liver and Spleen; Vector: pRT3D (Pharmacia)  
with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5', AACTGGAAGATTAATTAAGATCTTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pRT3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Patima Bonaldo."

FEATURES  
source  
1. .421  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3797717"  
/db\_xref="taxon:9606"  
/clone

DKFZp459D2446 5', mRNA sequence.  
 CR555382.1 GI:50248984  
 EST.  
 Pongo pygmaeus (orangutan)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.  
 REFERENCE  
 1 (bases 1 to 423)  
 Ansoerge, W., Krieger, S., Regiert, T., Rittmüller, C., Schwager, B.,  
 Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and  
 Wiemann, S.  
 TITLE  
 Pongo pygmaeus mRNA (Ansoerge, W., Krieger, S., Regiert, T., et al.)  
 JOURNAL  
 COMMENT  
 Unpublished (2004)  
 Contact: MIPS  
 MIPS  
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert from S. Wiemann,  
 Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
 Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL (European  
 Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA  
 sequencing consortium of the German Genome Project. This clone  
 (DKFZp459D2446) is available at the RZPD in Berlin. Please contact  
 the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
 information about the clone and the sequencing project is available  
 at http://mips.gsf.de/projects/cdna/.  
 Location/Qualifiers  
 1..423  
 /organism="Pongo pygmaeus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9600"  
 /clone="DKFZp459D2446"  
 /tissue type="cortex"  
 /dev stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="459 (synonym: pcorl)"  
 /note="vector: pSport1\_Sfi; Site\_1: SfiI; Site\_2: SfiIb"  
 ORIGIN  
 Query Match 93.3%; Score 14; DB 7; Length 423;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACACCTATGTGCA 14  
 |||||  
 Db 392 GACACCTATGTGCA 379  
 RESULT 36  
 BB815593  
 LOCUS  
 DEFINITION  
 BB815593 RIKEN full-length enriched, lung RCB-0558 LLC cDNA Mus  
 musculus cDNA clone G730029M05 3', mRNA sequence.  
 ACCESSION  
 BB815593.1 GI:16988222  
 VERSION  
 EST.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 424)  
 Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
 Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K.,  
 Iehii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,  
 Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,  
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,  
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akashi, S.,  
 Tanaka, T., Tomaru, A., Taya, T., Watahiki, A., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.  
 TITLE  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)

JOURNAL  
 COMMENT  
 Unpublished (2001)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
 and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-Format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
 Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.  
 Location/Qualifiers  
 1..424  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="G730029M05"  
 /tissue type="lung"  
 /cell\_line="RCB-0558 LLC"  
 /clone\_lib="RIKEN full-length enriched, lung RCB-0558 LLC  
 cDNA"  
 ORIGIN  
 Query Match 93.3%; Score 14; DB 2; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACACCTATGTGCA 14  
 |||||  
 Db 75 GACACCTATGTGCA 88  
 RESULT 37  
 AU018082/c  
 LOCUS  
 DEFINITION  
 AU018082 Mouse two-cell stage embryo cDNA Mus musculus cDNA clone  
 J0748F02 3', mRNA sequence.  
 ACCESSION  
 AU018082  
 VERSION  
 EST.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 425)  
 Ko, M.S.H., Kargul, J.R., Wang, X., Threat, T.A., Sun, T., DePalma, G.E.,  
 Liang, Y., Kitchin, G.J., Sharara, R. and Doi, H.  
 TITLE  
 Systematic analyses of genes expressed in 2-cell stage mouse  
 embryos (The ERATO/Doi Project at Wayne State University)  
 (Ko, M.S.H. et al.)  
 JOURNAL  
 COMMENT  
 Unpublished (1998)  
 Contact: Hirofumi Doi  
 Doi Bioasymmetry Project, ERATO  
 Japan Science and Technology Corporation (JST)  
 WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan

Unpublished (2001)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
 and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-Format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
 Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.  
 Location/Qualifiers  
 1..424  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="G730029M05"  
 /tissue type="lung"  
 /cell\_line="RCB-0558 LLC"  
 /clone\_lib="RIKEN full-length enriched, lung RCB-0558 LLC  
 cDNA"  
 ORIGIN  
 Query Match 93.3%; Score 14; DB 2; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACACCTATGTGCA 14  
 |||||  
 Db 75 GACACCTATGTGCA 88  
 RESULT 37  
 AU018082/c  
 LOCUS  
 DEFINITION  
 AU018082 Mouse two-cell stage embryo cDNA Mus musculus cDNA clone  
 J0748F02 3', mRNA sequence.  
 ACCESSION  
 AU018082  
 VERSION  
 EST.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 425)  
 Ko, M.S.H., Kargul, J.R., Wang, X., Threat, T.A., Sun, T., DePalma, G.E.,  
 Liang, Y., Kitchin, G.J., Sharara, R. and Doi, H.  
 TITLE  
 Systematic analyses of genes expressed in 2-cell stage mouse  
 embryos (The ERATO/Doi Project at Wayne State University)  
 (Ko, M.S.H. et al.)  
 JOURNAL  
 COMMENT  
 Unpublished (1998)  
 Contact: Hirofumi Doi  
 Doi Bioasymmetry Project, ERATO  
 Japan Science and Technology Corporation (JST)  
 WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan

Email: [hd@bioa.jst.go.jp](mailto:hd@bioa.jst.go.jp).

**FEATURES**  
source

Location/Qualifiers  
1..425  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="J0748F02"  
/dev\_stage="two-cell stage embryo"  
/clone\_lib="Mouse two-cell stage embryo cDNA"

**ORIGIN**

Query Match  
Best Local Similarity 93.3%; Score 14; DB 1; Length 425;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14  
|||||  
Db 96 GACACCTATGTGCA 83

**RESULT 38**  
AW471444  
LOCUS  
DEFINITION  
xw59e10.x1 NCI\_CGAP\_Pan1 Homo sapiens cDNA clone IMAGE:2832330 3', mRNA sequence.  
ACCESSION  
AW471444  
VERSION  
AW471444.1 GI:7041550  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM

**REFERENCE**  
AUTHORS  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
JOURNAL  
Tumor Gene Index  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Life Technologies catalog #: 11548-013  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -40UP from Gibco  
High quality sequence stop: 414.

**FEATURES**  
source

Location/Qualifiers  
1..426  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2832330"  
/tissue\_type="adenocarcinoma"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pan1"  
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo #11548-013"

**ORIGIN**

Query Match  
Best Local Similarity 93.3%; Score 14; DB 2; Length 426;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACCTATGTGCAC 15  
|||||  
Db 37 ACACCTATGTGCAC 50

**RESULT 39**  
BQ559024/c

**FEATURES**  
source

Location/Qualifiers  
1..427  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="niaEST:H4055F08-5"  
/db\_xref="taxon:10090"  
/clone="H4055F08"  
/sex="mixed"  
/dev\_stage="mixed"  
/lab\_host="DH10B"  
/clone\_lib="NIA Mouse 7.4K cDNA Clone Set"  
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 7,407 clones from more than 20 cDNA libraries."

**ORIGIN**

Query Match  
Best Local Similarity 93.3%; Score 14; DB 5; Length 427;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACCTATGTGCAC 15  
|||||  
Db 168 ACACCTATGTGCAC 155

**RESULT 40**  
BE074759/c

**LOCUS**  
BE074759  
DEFINITION  
IL5-BT0578-030400-034-e04 BT0578 Homo sapiens cDNA, mRNA sequence.  
ACCESSION  
BE074759  
VERSION  
BE074759.1 GI:8422978  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM

**REFERENCE**  
AUTHORS  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.

Email: [hd@bioa.jst.go.jp](mailto:hd@bioa.jst.go.jp).

**FEATURES**  
source

Location/Qualifiers  
1..425  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="J0748F02"  
/dev\_stage="two-cell stage embryo"  
/clone\_lib="Mouse two-cell stage embryo cDNA"

**ORIGIN**

Query Match  
Best Local Similarity 93.3%; Score 14; DB 1; Length 425;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 1 GACACCTATGTGCA 14  
|||||

**Db** 96 GACACCTATGTGCA 83

**RESULT 38**  
AW471444  
LOCUS  
DEFINITION  
xw59e10.x1 NCI\_CGAP\_Pan1 Homo sapiens cDNA clone IMAGE:2832330 3', mRNA sequence.  
ACCESSION  
AW471444  
VERSION  
AW471444.1 GI:7041550  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM

**REFERENCE**  
AUTHORS  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
JOURNAL  
Tumor Gene Index  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Life Technologies catalog #: 11548-013  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -40UP from Gibco  
High quality sequence stop: 414.

**FEATURES**  
source

Location/Qualifiers  
1..426  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2832330"  
/tissue\_type="adenocarcinoma"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pan1"  
/notes="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo #11548-013"

**ORIGIN**

Query Match  
Best Local Similarity 93.3%; Score 14; DB 2; Length 426;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 2 ACACCTATGTGCAC 15  
|||||

**Db** 37 ACACCTATGTGCAC 50

**RESULT 39**  
BQ559024/c

**FEATURES**  
source

Location/Qualifiers  
1..427  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="niaEST:H4055F08-5"  
/db\_xref="taxon:10090"  
/clone="H4055F08"  
/sex="mixed"  
/dev\_stage="mixed"  
/lab\_host="DH10B"  
/clone\_lib="NIA Mouse 7.4K cDNA Clone Set"  
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 7,407 clones from more than 20 cDNA libraries."

**ORIGIN**

Query Match  
Best Local Similarity 93.3%; Score 14; DB 5; Length 427;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 2 ACACCTATGTGCAC 15  
|||||

**Db** 168 ACACCTATGTGCAC 155

**RESULT 40**  
BE074759/c

**LOCUS**  
DEFINITION  
IL5-BT0578-030400-034-e04 BT0578 Homo sapiens cDNA, mRNA sequence.  
ACCESSION  
BE074759  
VERSION  
BE074759.1 GI:8422978  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM

**REFERENCE**  
AUTHORS  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

MEDLINE

PUBMED

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=IL5-BT0578-030>)

400-034-e04&t3=2000-04-03&t4=1

Seq primer: puc 18 forward

High quality sequence stop: 394.

#### FEATURES

source

1. .429

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="BT0578"

/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2:

SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No.

196,716 - Ludwig Institute for Cancer Research) profiles

into the puc 18 vector. Reverse transcription of tissue

mRNA and cDNA amplification were performed under low

stringency conditions."

#### ORIGIN

Query Match 93.3%; Score 14; DB 2; Length 429;

Best Local Similarity 100.0%; Pred. No. 3.6e+03;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACCTATGTGCA 14

Db 341 GACACCTATGTGCA 328

Search completed: June 4, 2005, 15:55:49

Job time : 768.212 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 10:28:02 ; Search time 1155.48 Seconds  
(without alignments)  
2138.683 Million cell updates/sec

Title: US-10-089-452-25  
Perfect score: 51  
Sequence: 1 agattgatccgcgaatgg.....atgacccgatattccagcc 51

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	51	100.0	51	6	AX113450 Sequence
2	51	100.0	51	6	AX113585 Sequence
3	51	100.0	51	6	AX816252 Sequence
4	51	100.0	360	6	AX113428 Sequence
5	51	100.0	360	6	AX113563 Sequence
6	51	100.0	360	6	AX816230 Sequence
7	44.6	87.5	312	10	MMU307489
8	44.6	87.5	312	10	MMU307503
9	43	84.3	192	10	MUSIGHCY
10	43	84.3	225	10	AF023210
11	43	84.3	234	10	MUSIGH2D
12	43	84.3	237	10	MUSIGHDA
13	43	84.3	248	10	MUSIGHDB
14	43	84.3	258	10	AY182513
15	43	84.3	273	10	AY172549
16	43	84.3	274	10	MMU307527
17	43	84.3	291	10	MUSIGHZN
18	43	84.3	294	10	MUSIGHZQ
19	43	84.3	294	10	MMU37861

20	43	84.3	294	10	MUSIGHZP	M33395	Mouse	Ig ge
21	43	84.3	297	10	AF062382	AF062382	Mus	muscu
22	43	84.3	300	10	AF546753	AF546753	Mus	muscu
23	43	84.3	306	10	MMU37847	U37847	Mus	musculu
24	43	84.3	309	10	AY182684	AY182684	Mus	muscu
25	43	84.3	309	10	AF059709	AF059709	Mus	muscu
26	43	84.3	311	10	MUSIGHJH	M17007	Mouse	Ig re
27	43	84.3	312	10	AY171989	AY171989	Mus	muscu
28	43	84.3	312	10	AY174032	AY174032	Mus	muscu
29	43	84.3	312	10	MMU307470	AJ307470	Mus	muscu
30	43	84.3	312	10	MMU307472	AJ307472	Mus	muscu
31	43	84.3	312	10	MMU307473	AJ307473	Mus	muscu
32	43	84.3	312	10	MMU307474	AJ307474	Mus	muscu
33	43	84.3	312	10	MMU307475	AJ307475	Mus	muscu
34	43	84.3	312	10	MMU307476	AJ307476	Mus	muscu
35	43	84.3	312	10	MMU307477	AJ307477	Mus	muscu
36	43	84.3	312	10	MMU307478	AJ307478	Mus	muscu
37	43	84.3	312	10	MMU307479	AJ307479	Mus	muscu
38	43	84.3	312	10	MMU307480	AJ307480	Mus	muscu
39	43	84.3	312	10	MMU307481	AJ307481	Mus	muscu
40	43	84.3	312	10	MMU307482	AJ307482	Mus	muscu
41	43	84.3	312	10	MMU307484	AJ307484	Mus	muscu
42	43	84.3	312	10	MMU307485	AJ307485	Mus	muscu
43	43	84.3	312	10	MMU307486	AJ307486	Mus	muscu
44	43	84.3	312	10	MMU307487	AJ307487	Mus	muscu
45	43	84.3	312	10	MMU307488	AJ307488	Mus	muscu

#### ALIGNMENTS

RESULT 1  
AX113450  
LOCUS  
DEFINITION Sequence 25 from Patent WO0127612.  
ACCESSION AX113450  
VERSION AX113450.1 GI:13939709  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Reiter,C., Cullmann,G., Lakner,M., Truee,A., Dehnert,S. and Schwartz,G.  
TITLE Immuno-chromatographic rapid assay in order to detect acid-resistant microorganisms in the stool  
JOURNAL Patent: WO 0127612-A 25 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung mbH (DE)

51 bp DNA linear PAT 01-MAY-2001

FEATURES  
source Location/Qualifiers

1..51  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="CDR"

#### ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 51;  
Best Local Similarity 100.0%; Pred. No. 3.3e-08;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
|||||  
Db 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51

#### RESULT 2

AX113585  
LOCUS  
DEFINITION Sequence 25 from Patent WO0127613.  
ACCESSION AX113585  
VERSION AX113585.1 GI:13939780  
KEYWORDS

51 bp DNA linear PAT 01-MAY-2001

SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.

1  
REFERENCE  
AUTHORS  
Reiter,C., Cullmann,G., Heppner,P., Ringeis,A., Mueller,H. and  
Haindl,E.

TITLE  
Improved method for the detection of acid resistant microorganisms  
in a stool

JOURNAL  
Patent: WO 0127613-A 25 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung  
(DE)

FEATURES  
source  
Location/Qualifiers  
1..51  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="CDR"

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 51; DB 6; Length 51;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGAATGCTAAACTAAATATGACCCGATATTCAGGCC 51  
|||||  
Db 1 AAGATTGATCCTCGAATGCTAAACTAAATATGACCCGATATTCAGGCC 51  
|||||

RESULT 3  
AX816252  
LOCUS  
AX816252  
DEFINITION  
Sequence 25 from Patent EP1336850.  
ACCESSION  
AX816252  
VERSION  
AX816252.1 GI:39646774  
KEYWORDS  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.

1  
REFERENCE  
AUTHORS  
Reiter,C., Cullmann,G., Mueller,H., Heppner,P., Haindl,E. and  
Ringeis,A.

TITLE  
Improved method for the detection of acid resistant microorganisms  
in a stool

JOURNAL  
Patent: EP 1336850-A 25 20-AUG-2003;  
Connex Gesellschaft zur Optimierung von Forschung und Ent wicklung  
(DE)

FEATURES  
source  
Location/Qualifiers  
1..51  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 51; DB 6; Length 51;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGAATGCTAAACTAAATATGACCCGATATTCAGGCC 51  
|||||  
Db 1 AAGATTGATCCTCGAATGCTAAACTAAATATGACCCGATATTCAGGCC 51  
|||||

RESULT 4  
AX113428  
LOCUS  
AX113428  
DEFINITION  
Sequence 3 from Patent WO0127612.  
ACCESSION  
AX113428  
VERSION  
AX113428.1 GI:13939696  
KEYWORDS  
Mus musculus (house mouse)  
Mus musculus

1  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
Reiter,C., Cullmann,G., Lakner,M., Truee,A., Dehnert,S. and  
Schwartz,G.

TITLE  
Immuno-chromatographic rapid assay in order to detect  
acid-resistant microorganisms in the stool

JOURNAL  
Patent: WO 0127612-A 3 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung  
mbH (DE)

FEATURES  
source  
Location/Qualifiers  
1..360  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 51; DB 6; Length 360;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGAATGCTAAACTAAATATGACCCGATATTCAGGCC 51  
|||||  
Db 148 AAGATTGATCCTCGAATGCTAAACTAAATATGACCCGATATTCAGGCC 198  
|||||

RESULT 5  
AX113563  
LOCUS  
AX113563  
DEFINITION  
Sequence 3 from Patent WO0127613.  
ACCESSION  
AX113563  
VERSION  
AX113563.1 GI:13939767  
KEYWORDS  
Mus musculus (house mouse)  
Mus musculus

1  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1  
REFERENCE  
AUTHORS  
Reiter,C., Cullmann,G., Heppner,P., Ringeis,A., Mueller,H. and  
Haindl,E.

TITLE  
Improved method for the detection of acid resistant microorganisms  
in a stool

JOURNAL  
Patent: WO 0127613-A 3 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung  
(DE)

FEATURES  
source  
Location/Qualifiers  
1..360  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"

ORIGIN

Query Match  
Best Local Similarity 100.0%; Score 51; DB 6; Length 360;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGAATGCTAAACTAAATATGACCCGATATTCAGGCC 51  
|||||  
Db 148 AAGATTGATCCTCGAATGCTAAACTAAATATGACCCGATATTCAGGCC 198  
|||||

RESULT 6  
AX816230  
LOCUS  
AX816230  
DEFINITION  
Sequence 3 from Patent EP1336850.  
ACCESSION  
AX816230  
VERSION  
AX816230.1 GI:39646761  
KEYWORDS  
Mus musculus (house mouse)  
Mus musculus

1  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1  
REFERENCE  
AUTHORS  
Reiter,C., Cullmann,G., Mueller,H., Heppner,P., Haindl,E. and  
Ringeis,A.



**TITLE** Improved method for the detection of acid resistant microorganisms in a stool  
**JOURNAL** Patent: EP 1336850-A 3 20-AUG-2003;  
Connex Gesellschaft zur Optimierung von Forschung und Ent wicklung (DE)

**FEATURES**  
source Location/Qualifiers  
1. .360  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"

**ORIGIN**  
Query Match 100.0%; Score 51; DB 6; Length 360;  
Best Local Similarity 100.0%; Pred. No. 3.2e-08;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCGGATATCCAGGCC 51  
|||||  
Db 148 AAGATTGATCTCGGAATGGTAAACTAAATATGACCGGATATCCAGGCC 198  
|||||

**RESULT 7**  
MMU307489  
LOCUS Mus musculus partial 17.2.25-D-Jh4 DNA linear ROD 07-FEB-2002  
DEFINITION chain, sequence t416.  
ACCESSION AJ307489  
VERSION 17.2.25-D-Jh4 gene; immunoglobulin heavy chain.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

**REFERENCE**  
1 Toellner K.M., Jenkinson, W.E., Taylor, D.R., Khan, M., Sze, D.M., Sansom, D.M., Vinuesa, C.G. and MacLennan, I.C.  
Low-level hypermutation in T cell-independent germinal centers compared with high mutation rates associated with T cell-dependent germinal centers  
J. Exp. Med. 195 (3), 383-389 (2002)  
21686239  
PUBMED 11828014  
Toellner, K.M.  
Direct Submission  
Submitted (01-FEB-2001) Toellner K.M., Immunology, University of Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED KINGDOM

**FEATURES**  
source Location/Qualifiers  
1. .312  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/cell\_type="germinal center"  
/tissue\_type="spleen"  
/rearranged  
/note="4 d after NP/CGG immunisation in CGG primed recipients of QM B cells"  
1. .312  
/gene="17.2.25-D-Jh4"  
<1. .>312  
/gene="17.2.25-D-Jh4"  
/codon\_start=1  
/product="immunoglobulin heavy chain"  
/protein\_id="CAC29292.1"  
/db\_xref="GI:12964108"  
/translation="GAEIVKPGASVKLSCTASGPNIKDTYMHVWVKRPEQGLEWIGRI DPANGTKYDPKFGKATITADTSSNTAYLQSLTSEDYAVYCYRYRPPYAMDY WG"

**gene** 1. .273  
/gene="17.2.25"  
<1. .>273  
/gene="17.2.25"

**CDS**  
1. .312  
/gene="17.2.25-D-Jh4"  
<1. .>312  
/gene="17.2.25-D-Jh4"  
/codon\_start=1  
/product="immunoglobulin heavy chain"  
/protein\_id="CAC29292.1"  
/db\_xref="GI:12964108"  
/translation="GAEIVKPGASVKLSCTASGPNIKDTYMHVWVKRPEQGLEWIGRI DPANGTKYDPKFGKATITADTSSNTAYLQSLTSEDYAVYCYRYRPPYAMDY WG"

**gene** 1. .273  
/gene="17.2.25"  
<1. .>273  
/gene="17.2.25"

**V\_segment**  
1. .273  
/gene="17.2.25"

**gene** 288. .312  
/gene="Jh4"  
288. .>312  
/gene="Jh4"

**J\_segment**  
288. .>312  
/gene="Jh4"

**ORIGIN**  
Query Match 87.5%; Score 44.6; DB 10; Length 312;  
Best Local Similarity 92.2%; Pred. No. 7.6e-06;  
Matches 47; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCGGATATCCAGGCC 51  
|||||  
Db 127 AGGATTGATCTCGGAATGGTAAACTAAATATGACCGGATATCCAGGCC 177  
|||||

**RESULT 8**  
MMU307503  
LOCUS Mus musculus partial 17.2.25-D-Jh4 DNA linear ROD 07-FEB-2002  
DEFINITION chain, sequence x31c.  
ACCESSION AJ307503  
VERSION 17.2.25-D-Jh4 gene; immunoglobulin heavy chain.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

**REFERENCE**  
1 Toellner K.M., Jenkinson, W.E., Taylor, D.R., Khan, M., Sze, D.M., Sansom, D.M., Vinuesa, C.G. and MacLennan, I.C.  
Low-level hypermutation in T cell-independent germinal centers compared with high mutation rates associated with T cell-dependent germinal centers  
J. Exp. Med. 195 (3), 383-389 (2002)  
21686239  
PUBMED 11828014  
Toellner, K.M.  
Direct Submission  
Submitted (01-FEB-2001) Toellner K.M., Immunology, University of Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED KINGDOM

**FEATURES**  
source Location/Qualifiers  
1. .312  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/cell\_type="germinal center"  
/tissue\_type="spleen"  
/rearranged  
/note="4 d after NP-Ficoll immunisation of QM B cell recipients"  
1. .312  
/gene="17.2.25-D-Jh4"  
<1. .>312  
/gene="17.2.25-D-Jh4"  
/codon\_start=1  
/product="immunoglobulin heavy chain"  
/protein\_id="CAC29306.1"  
/db\_xref="GI:12964136"  
/translation="GAEIVKPGASVKLSCTASGPNIKDTYMHVWVKRPEQGLEWIGRI DPANGTKYDPKFGKATITADTSSNTAYLQSLTSEDYAVYCYRYRPPYAMDY WG"

**gene** 1. .273  
/gene="17.2.25"  
<1. .>273  
/gene="17.2.25"

**CDS**  
1. .312  
/gene="17.2.25-D-Jh4"  
<1. .>312  
/gene="17.2.25-D-Jh4"  
/codon\_start=1  
/product="immunoglobulin heavy chain"  
/protein\_id="CAC29306.1"  
/db\_xref="GI:12964136"  
/translation="GAEIVKPGASVKLSCTASGPNIKDTYMHVWVKRPEQGLEWIGRI DPANGTKYDPKFGKATITADTSSNTAYLQSLTSEDYAVYCYRYRPPYAMDY WG"

**gene** 1. .273  
/gene="17.2.25"  
<1. .>273  
/gene="17.2.25"

**V\_segment**  
1. .273  
/gene="17.2.25"

**ORIGIN**  
1. .273  
/gene="17.2.25"  
<1. .>273  
/gene="17.2.25"

Query Match	87.5%;	Score 44.6;	DB 10;	Length 312;
Best Local Similarity	92.2%;	Pred. No. 7.6e-06;		
Matches	47;	Conservative 0;	Mismatches 4;	Indels 0; Gaps 0;
Qy	1	AAGATTGATCTCGCAATGTTAAACTAAATATGACCCGATATTCAGGCC	51	
Db	127	AGGATTGATCTCGCAATGTTAACTAAATATGACCCGAATATTCAGGCC	177	
RESULT 9				
MUSIGHCY				
LOCUS	MUSIGHCY	192 bp	mRNA	linear
DEFINITION	Mouse Ig active mu-chain anti-GAT VDJ4-region mRNA	from HP20.33.		
ACCESSION	M13066			
VERSION	M13066.1	GI:195192		
KEYWORDS	C-region; D-region; J-region; V-region; immunoglobulin heavy chain; immunoglobulin mu-chain; processed gene.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 192)			
AUTHORS	Roth,C., Rocca-Serra,J., Somee,G., Fougereau,M. and Theze,J.			
TITLE	Gene repertoire of the anti-poly(Glu6Ala30Tyr10) (GAT) immune response: comparison of VH, V kappa, and D regions used by anti-GAT antibodies and monoclonal antibodies produced after anti-idiotypic immunization			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 82 (14), 4788-4792 (1985)			
MEDLINE	85270412			
PUBMED	327291			
COMMENT	Original source text: Mouse (Balb/c) monoclonal antibody 20.33, cDNA to mRNA, generated after anti-idiotypic immunization with HP-Id20-KLH.			
FEATURES	Draft entry and clean copy of the sequence in [1] were kindly provided by C.Roth, 30-SEP-1985.			
source	Location/Qualifiers			
	1..192			
	/organism="Mus musculus"			
	/mol_type="mRNA"			
	/db_xref="taxon:10090"			
	<1..>192			
	/notes="Ig mu-chain (VDJ4)"			
	/codon_start=1			
	/protein_id="AAA38194.1"			
	/db_xref="GI:195193"			
	/translation="GRIDPANGTKYDPKFGKATITADTSSNTAVLQSLTSDTA VYCARSSSTSYAMDYWGQGG"			
ORIGIN	Chromosome 12.			
Query Match	84.3%;	Score 43;	DB 10;	Length 192;
Best Local Similarity	90.2%;	Pred. No. 3e-05;		
Matches	46;	Conservative 0;	Mismatches 5;	Indels 0; Gaps 0;
Qy	1	AAGATTGATCTCGCAATGTTAAACTAAATATGACCCGATATTCAGGCC	51	
Db	4	AGGATTGATCTCGCAATGTTAACTAAATATGACCCGAATATTCAGGCC	54	
RESULT 10				
AF023210				
LOCUS	AF023210	225 bp	mRNA	linear
DEFINITION	Mus musculus clone HV3-P3G, family J558 anti-fluorescein monoclonal IgM heavy chain mRNA, partial cds.			
ACCESSION	AF023210			
VERSION	AF023210.1	GI:2570546		
KEYWORDS				
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 225)			
AUTHORS	Van der Key,H., Hsu,C., Tolat,A., Kansil,S., Dalesandro,M.R., Dorer,D.R., Caton,A. and Owen,J.A.			

```
CDS
<1..>234
/note="Ig H-chain V-D-JH4-region"
/codon_start=1
/protein_id="AAA38654.1"
/db_xref="GI:196294"
/translation="OGXKWIGRIDPANGNTKYDPKFGKATITADTSSNTAYLQLSSL
TSETAVYCCARRCGNGYANDYWGQSTVTSV"
ORIGIN
Chromosome 12.

Query Match      84.3%; Score 43; DB 10; Length 234;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
|
Db 22 AGGATTGATCTCGAATGGTAAATACTAAATATGACCGAAGTTCAGGCC 72
|

RESULT 12
MUSIGHDA
LOCUS
DEFINITION
Mouse Ig active mu-chain anti-GAT VDJ4-region mRNA from HP22.8.
ACCESSION
M13068
VERSION
GI:195198
KEYWORDS
C-region; D-region; J-region; V-region; immunoglobulin heavy chain;
immunoglobulin mu-chain; processed gene.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 237)
/protein_id="AAA38196.1"
/db_xref="GI:195199"
/translation="IKDTYHWVKQRPQGLEWIGRIDPANGNTKYDPKFGKATITA
DTSNTAYLQLSSLTSDTYVYCARWESYWGQG"
Chromosome 12.

Query Match      84.3%; Score 43; DB 10; Length 237;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
|
Db 64 AGGATTGATCTCGAATGGTAAATACTAAATATGACCGAAGTTCAGGCC 114
|

RESULT 13
MUSIGHDB
LOCUS
DEFINITION
Mouse Ig active mu-chain anti-GAT VDJ4-region mRNA from HP22.176.
ACCESSION
M13069
VERSION
GI:195200
```

```
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 248)
Roth,C., Rocca-Serra,J., Somme,G., Fougereau,M. and These,J.
Gene repertoire of the anti-poly(Glu60Ala30Tyr10) (GAT) immune
response: comparison of VH, V kappa, and D regions used by anti-GAT
antibodies and monoclonal antibodies produced after anti-idiotypic
immunization
Proc. Natl. Acad. Sci. U.S.A. 82 (14), 4788-4792 (1985)
85270412
PUBMED
3927291
COMMENT
Original source text: Mouse (Balb/c) monoclonal antibody 22.176,
HP-Id22-KLH,
generated after anti-idiotypic immunization with
Draft entry and clean copy of the sequence in [1] were kindly
provided by C.Roth, 30-SEP-1985.
FEATURES
source
1..248
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
<1..>248
/note="Ig mu-chain (VDJ4)"
/codon_start=1
/protein_id="AAA38197.1"
/db_xref="GI:553983"
/translation="IKDTYHWVKQRPQGLEWIGRIDPANGNTKYDPKFGKATITA
DTSNTAYLQLSSLTSDTYVYCARSDGYLLWTGVK"
Chromosome 12.

Query Match      84.3%; Score 43; DB 10; Length 248;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
|
Db 64 AGGATTGATCTCGAATGGTAAATACTAAATATGACCGAAGTTCAGGCC 114
|

RESULT 14
AY182513
LOCUS
DEFINITION
Mus musculus clone BaFl-P40 immunoglobulin heavy chain variable
region mRNA, partial cds.
ACCESSION
AY182513
VERSION
AY182513.1
KEYWORDS
GI:27728817
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 258)
Kretschmer,K., Engel,H. and Weiss,S.
Strong antigenic selection shaping the immunoglobulin heavy chain
repertoire of B-1a lymphocytes in lambda 2(315) transgenic mice
Eur. J. Immunol. 32 (8), 2317-2327 (2002)
23197943
PUBMED
12209645
REFERENCE
2 (bases 1 to 258)
Kretschmer,K., Engel,H. and Weiss,S.
Direct Submission
Submitted (18-NOV-2002) Molecular Immunology, GBF, German Research
Centre for Biotechnology, Mascheroder Weg 1, Braunschweig 38124,
Germany
FEATURES
source
1..258
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
```

```

/db_xref="taxon:10090"
/clonetype="B lymphocyte"
/tissue_type="liver"
/dev stage="fetus"
/notes="lambda2 MOPC315 transgenic L2 mouse"
<1..>258
/codon_start=1
/product="immunoglobulin heavy chain variable region"
/protein_id="AA018869.1"
/db_xref="GI:27728818"
/translation="YHMKVQRPQGLEWIGRIDPANGNTKYPKFGKATITADTSS
NTAYLQLSSLTSEDYAVYCARSTMITTFAYWQGLTVTVA"

CDs
Query Match      84.3%; Score 43; DB 10; Length 258;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy  1  AGATTGATCCTCGAATGTAATAAATAATGACCCGATATTCAGGCC 51
|  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  52  AGGATTGATCCTCGAATGTAATAAATAATGACCCGATATTCAGGCC 102
|  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 15
LOCUS      AY172549                273 bp      mRNA      linear      ROD 22-JAN-2003
DEFINITION Mus musculus clone L2pCD5negP42 immunoglobulin heavy chain
SOURCE      variable region mRNA, partial cds.
ACCESSION   AY172549
VERSION     AY172549.1 GI:27818382
KEYWORDS    Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 273)
AUTHORS    Kretschmer,K., Jungbloud,A., Stopkowicz,J., Hoffmann,R. and
            Weiss,S.
TITLE      Antibody repertoire and gene expression profile of splenic and
            peritoneal B-1a lymphocytes: implications on different
            developmental and functional traits
JOURNAL
REFERENCE   2  (bases 1 to 273)
AUTHORS    Kretschmer,K., Jungbloud,A., Stopkowicz,J., Hoffmann,R. and
            Weiss,S.
TITLE      Direct Submission
JOURNAL    Submitted (04-NOV-2002) Molecular Immunology, GBF, German Research
            Centre for Biotechnology, Mascheroder Weg 1, Braunschweig 38124,
            Germany

FEATURES             Location/Qualifiers
     source           1..273
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="BA1B/c"
                     /db_xref="taxon:10090"
                     /clone="L2pCD5negP42"
                     /clonetype="B lymphocyte"
                     /tissue_type="spleen"
                     /dev_stage="adult"
                     /notes="lambda2 MOPC315 transgenic L2 mouse"
                     <1..>273
                     /codon_start=1
                     /product="immunoglobulin heavy chain variable region"
                     /protein_id="AA023274.1"
                     /db_xref="GI:27818383"
                     /translation="KDTYHMKVQRPQGLEWIGRIDPANGNTKYPKFGKATITAD
                     TSSNTAYLQLSSLTSEDYAVYCAPDYGTGTFAYWQGLTVTVA"

ORIGIN
Query Match      84.3%; Score 43; DB 10; Length 273;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

CDs
Query Match      84.3%; Score 43; DB 10; Length 258;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy  1  AGATTGATCCTCGAATGTAATAAATAATGACCCGATATTCAGGCC 51
|  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  52  AGGATTGATCCTCGAATGTAATAAATAATGACCCGATATTCAGGCC 102
|  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 16
LOCUS      MMU307527                274 bp      DNA      linear      ROD 10-MAR-2001
DEFINITION Mus musculus partial 17.2.25-D-Jh4 gene for immunoglobulin heavy
SOURCE      chain, sequence f43a.
ACCESSION   AJ307527.1 GI:12964183
VERSION     AJ307527.1 GI:12964183
KEYWORDS    Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 274)
AUTHORS    Toellner,K.M., Jenkinson,W.E., Garcia de Vinuesa,C., Taylor,D.R.
            and MacLennan,I.C.M.
TITLE      Lack of hypermutation in T cell independent germinal centers
JOURNAL
REFERENCE   2  (bases 1 to 274)
AUTHORS    Toellner,K.M.
TITLE      Direct Submission
JOURNAL    Submitted (01-FEB-2001) Toellner K.M., Immunology, University of
            Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED
            KINGDOM

FEATURES             Location/Qualifiers
     source           1..274
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /cell_type="germinal center"
                     /tissue_type="spleen"
                     /rearranged
                     /notes="4 d after NP-Ficoll immunisation of QM mice"
                     1..96
                     /genes="17.2.25"
                     1..96
                     /genes="17.2.25-D-Jh4"
                     <1..96
                     /genes="17.2.25-D-Jh4"
                     /codon_start=1
                     /product="immunoglobulin heavy chain"
                     /protein_id="CAC29330.1"
                     /db_xref="GI:12964184"
                     /translation="GAELVPGASVKLSCTASGFNIKPGVDWKD"
                     <1..96
                     /genes="17.2.25"

ORIGIN
Query Match      84.3%; Score 43; DB 10; Length 274;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy  1  AGATTGATCCTCGAATGTAATAAATAATGACCCGATATTCAGGCC 51
|  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  89  AGGATTGATCCTCGAATGTAATAAATAATGACCCGATATTCAGGCC 139
|  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 17
LOCUS      MUSIGHZN                291 bp      DNA      linear      ROD 27-APR-1993
DEFINITION Mouse Ig germline H-chain gene V-region, Id+ hybridoma 41-4F5,
SOURCE      partial cds.
ACCESSION   M33393.1 GI:196309
VERSION     M33393.1 GI:196309
KEYWORDS    V-region; germline; immunoglobulin heavy chain.
SOURCE      Mus musculus (house mouse)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 291)
REFERENCE
AUTHORS Borriero,L., Giorgetti,C., Smith,G., Landry,D., Selsing,E.,
Zhukovsky,E. and Press,J.L.
TITLE Neonatal and adult primary B cells use the same germ-line VH and V
kappa genes in their (T,G)-A-L-specific repertoire
J. Immunol. 144 (2), 583-592 (1990)
JOURNAL
MEDLINE 90111072
PUBMED 2129539
COMMENT Original source text: Mouse (strain Balb.b) neonatal Id+ hybridoma
41-4F5 DNA.
FEATURES
source Location/Qualifiers
1..291
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>291
/note="Ig H-chain V-region"
/codon_start=1
/protein_id="AAA38657.1"
/db_xref="GI:196310"
/translation="EVQLQSGXELVPGASVKLSCTASGFINIKDTYMHVWKORPEQG
LEWIGRIDPANGNTKYDPKFGKATITADTSSNTAYLQLSLTSEDVAVYCA"
ORIGIN
Chromosome 12.
Query Match 84.3%; Score 43; DB 10; Length 291;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGAATGCTGTAATAAATAATGACCGCATATTCAGGGCC 51
Db 148 AGGATTGATCTCGAATGCTGTAATAAATAATGACCGCATATTCAGGGCC 198

RESULT 18
MUSIGHZQ 291 bp DNA linear ROD 27-APR-1993
LOCUS Mouse Ig germline H-chain gene V-region, Id+ hybridoma 44-2F5,
DEFINITION partial cds.
VERSION M33396.1 GI:196315
KEYWORDS V-region; germline; immunoglobulin heavy chain.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 291)
REFERENCE
AUTHORS Borriero,L., Giorgetti,C., Smith,G., Landry,D., Selsing,E.,
Zhukovsky,E. and Press,J.L.
TITLE Neonatal and adult primary B cells use the same germ-line VH and V
kappa genes in their (T,G)-A-L-specific repertoire
J. Immunol. 144 (2), 583-592 (1990)
JOURNAL
MEDLINE 90111072
PUBMED 2129539
COMMENT Original source text: Mouse (strain Balb.b) neonatal Id+ hybridoma
44-2F5 DNA.
FEATURES
source Location/Qualifiers
1..291
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>291
/note="Ig H-chain V-region"
/codon_start=1
/protein_id="AAA38660.1"
/db_xref="GI:196316"
/translation="EVQLQSGXELVPGASVKLSCTASGFINIKDTYMHVWKORPEQG
LEWIGRIDPANGNTKYDPKFGKATITADTSSNTAYLQLSLTSEDVAVYCA"
ORIGIN
Chromosome 12.
Query Match 84.3%; Score 43; DB 10; Length 291;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGAATGCTGTAATAAATAATGACCGCATATTCAGGGCC 51
Db 148 AGGATTGATCTCGAATGCTGTAATAAATAATGACCGCATATTCAGGGCC 198

RESULT 19
MUSIGHZP 294 bp DNA linear ROD 27-APR-1993
LOCUS Mouse Ig germline H-chain gene V-region, Id+ hybridoma 23-10E7,
DEFINITION partial cds.
VERSION M33395.1 GI:196313
KEYWORDS V-region; germline; immunoglobulin heavy chain.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 294)
REFERENCE
AUTHORS Borriero,L., Giorgetti,C., Smith,G., Landry,D., Selsing,E.,

```

```

Qy 1 AAGATTGATCTCGAATGCTGTAATAAATAATGACCGCATATTCAGGGCC 51
Db 148 AGGATTGATCTCGAATGCTGTAATAAATAATGACCGCATATTCAGGGCC 198

RESULT 19
MUSIGHZP 294 bp DNA linear ROD 27-APR-1993
LOCUS Mouse Ig germline H-chain gene V-region, Id+ hybridoma 23-10E7,
DEFINITION partial cds.
VERSION M33395.1 GI:196313
KEYWORDS V-region; germline; immunoglobulin heavy chain.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 294)
REFERENCE
AUTHORS Borriero,L., Giorgetti,C., Smith,G., Landry,D., Selsing,E.,

```

Zukowsky,E. and Press,J.L.  
Neonatal and adult primary B cells use the same germ-line VH and V kappa genes in their (T,G)-A-L-specific repertoire  
J. Immunol. 144 (2), 583-592 (1990)  
90111072  
PUBMED 2129539  
COMMENT Original source text: Mouse (strain Balb.b) neonatal Id+ hybridoma 23-1057 DNA.

FEATURES  
source Location/Qualifiers  
1..294  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
<1..>294  
/note="Ig H-chain V-region"  
/codon\_start=1  
/protein\_id="AAA38659.1"  
/db\_xref="GI:196314"  
/translation="EVQLQSGAEIVKPGASVKLSCTASGXNFKDTYMHVWKORPEGL  
LEWIGRIDPANGNTKYDPKFGKATITADTSSNTAVTLQLSLTSEDYAVYYCAR"  
Chromosome 12.

ORIGIN  
Query Match 84.3%; Score 43; DB 10; Length 294;  
Best Local Similarity 90.2%; Pred.No. 3e-05;  
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGGAATGGTTAAACATAATATGCCCGCATATTCACAGGC 51  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 148 AGATTGATCTCGGAATGGTTACTTAATATGACCAGAATTCACAGGC 198

RESULT 21  
AF062382 AF062382 297 bp DNA linear ROD 01-MAY-2001  
LOCUS Mus musculus clone A2.a immunoglobulin heavy chain VDJ region gene,  
partial cds.  
AF062382  
AF062382.1 GI:3818534

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1..297  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="BALB.K"  
/db\_xref="taxon:10090"  
/clone="A2.a"  
/dev stage="adult"  
/rearranged  
/note="Obtained fourteen days after priming with the antigen Np-CGG"  
<1..>297  
/product="immunoglobulin heavy chain VDJ region"  
<1..>297  
/codon\_start=1  
/product="immunoglobulin heavy chain VDJ region"  
/protein\_id="AAC69553.1"  
/db\_xref="GI:3818535"  
/translations="LQQSGAEIVKPGASVKLSCTAAGFNFKDTYIHVWKORPEGLE"

```

ACCESSION      U37847
VERSION        U37847.1  GI:1127608
KEYWORDS
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 306)
AUTHORS        Caton,A.J., Swartzentruber,J.R., Kuhl,A.L., Carding,S.R. and
               Stark,S.B.
TITLE          Activation and negative selection of functionally distinct subsets
               of antibody-secreting cells by influenza hemagglutinin as a viral
               and a neo-self antigen
               J. Exp. Med. 183 (1), 13-26 (1996)
JOURNAL        96136744
MEDLINE        8551216
PUBMED
REFERENCE      2 (bases 1 to 306)
AUTHORS        Swartzentruber,J.R.
TITLE          Direct Submission
JOURNAL        Submitted (10-OCT-1995) Jennifer R. Swartzentruber, The Wistar
               Institute, 3601 Spruce St., Philadelphia, PA 19104, USA
FEATURES       Location/Qualifiers
               1..306
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="BALB/c"
               /isolate="h1343-84"
               /db_xref="taxon:10090"
               /cell_type="B-lymphocyte"
               <1..>306
               /codon_start=1
               /product="rearranged immunoglobulin heavy chain variable
               region"
               /protein_id="AAA92370.1"
               /db_xref="GI:1127609"
               /translation="KLCTASGFNIKDTYMHVVKORPEQGLEWIGRIDPANGNTKYDP
               KPQGKATITADTSSNTAYLQLSLTSEDVAVYCYCRYPDYWGQGTLLTVS
               "
CDS
               1..306
               /codon_start=1
               /product="rearranged immunoglobulin heavy chain variable
               region"
               /protein_id="AAA92370.1"
               /db_xref="GI:1127609"
               /translation="KLCTASGFNIKDTYMHVVKORPEQGLEWIGRIDPANGNTKYDP
               KPQGKATITADTSSNTAYLQLSLTSEDVAVYCYCRYPDYWGQGTLLTVS
               "
ORIGIN
               Query Match      84.3%; Score 43; DB 10; Length 306;
               Best Local Similarity 90.2%; Pred. No. 3e-05;
               Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 AAGATTGATCTCGGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
    |||||
Db 94 AGGATTGATCTCGGAATGTTAAACTAAATATGACCCGAGTTCAGGCC 144
    |||||

RESULT 24
AY182684
LOCUS          AY182684
DEFINITION    Mus musculus clone L2F1ld-P3 immunoglobulin heavy chain variable
               region mRNA, partial cds.
ACCESSION     AY182684
VERSION       AY182684.1  GI:27729497
KEYWORDS
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 309)
AUTHORS        Kretscher,K., Engel,H. and Weiss,S.
TITLE          Strong antigenic selection shaping the immunoglobulin heavy chain
               repertoire of B-1a lymphocytes in lambda 2(315) transgenic mice
               Eur. J. Immunol. 32 (8), 2317-2327 (2002)
JOURNAL        22197943
MEDLINE        12209645
PUBMED
REFERENCE      2 (bases 1 to 309)
AUTHORS        Kretscher,K., Engel,H. and Weiss,S.
TITLE          Direct Submission
JOURNAL        Submitted (18-NOV-2002) Molecular Immunology, GBF, German Research
               Centre for Biotechnology, Mascheroder Weg 1, Braunschweig 38124,

```

```

Germany
FEATURES       source
               Location/Qualifiers
               1..309
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="BALB/c"
               /isolation_source="in vitro differentiated liver"
               /db_xref="taxon:10090"
               /clone="L2F1ld-P3"
               /cell_type="B lymphocyte"
               /dev_stage="fetus"
               /notes="lambda2 MOPC315 transgenic L2 mouse"
               <1..>309
               /codon_start=1
               /product="immunoglobulin heavy chain variable region"
               /protein_id="AA019040.1"
               /db_xref="GI:27729498"
               /translation="LVKPGASVKLSCTASGFNIKDTYMHVVKORPEQGLEWIGRIDPA
               NGNTKYDPKFGKATITADTSSNTAYLQLSLTSEDVAVYCYCRYPDYWGQGTLLTVS
               S"
CDS
               1..309
               /codon_start=1
               /product="immunoglobulin heavy chain variable region"
               /protein_id="AA019040.1"
               /db_xref="GI:27729498"
               /translation="LVKPGASVKLSCTASGFNIKDTYMHVVKORPEQGLEWIGRIDPA
               NGNTKYDPKFGKATITADTSSNTAYLQLSLTSEDVAVYCYCRYPDYWGQGTLLTVS
               S"
ORIGIN
               Query Match      84.3%; Score 43; DB 10; Length 309;
               Best Local Similarity 90.2%; Pred. No. 3e-05;
               Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 AAGATTGATCTCGGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
    |||||
Db 118 AGGATTGATCTCGGAATGTTAAACTAAATATGACCCGAGTTCAGGCC 168
    |||||

RESULT 25
AF059709
LOCUS          AF059709
DEFINITION    Mus musculus clone N2.3.a immunoglobulin heavy chain VDJ region
               mRNA, partial cds.
ACCESSION     AF059709
VERSION       AF059709.1  GI:3859461
KEYWORDS
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 309)
AUTHORS        Gorgetti,C.A. and Press,J.L.
TITLE          Somatic mutation in the neonatal mouse
JOURNAL        J. Immunol. 161 (11), 6093-6104 (1998)
MEDLINE        95049839
PUBMED
REFERENCE      2 (bases 1 to 309)
AUTHORS        Press,J.L.
TITLE          Direct Submission
JOURNAL        Submitted (16-APR-1998) Rosenstiel Center, Brandeis University, 415
               South St, Waltham, MA 02254-9110, USA
FEATURES       Location/Qualifiers
               1..309
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="BALB.K"
               /db_xref="taxon:10090"
               /clone="N2.3.a"
               /dev_stage="one-day old neonate"
               /rearranged
               /notes="RT-PCR product; obtained fourteen days after
               priming with the antigen Np-CGG"
               <1..>309
               /codon_start=1
               /product="immunoglobulin heavy chain VDJ region"
               /protein_id="AAC72733.1"
               /db_xref="GI:3859462"
               /translation="EVLQSQGAELVKPGASVKLSCTASGFNIKDTYMHVVKORPEQ
               LEWIGRIDPANGNTKYDPKFGKATITADTSSNTAYLQLSLTSEDVAVYCYCAKGY
               Y"

```





```

/clone="BapB1-46"
/cell_type="B-1a lymphocyte"
/tissue_type="peritoneal cavity"
/dev_stage="adult"
<1..>312
/codon_start=1
/product="immunoglobulin heavy chain variable region"
/protein_id="AA019722.1"
/db_xref="GI:27752518"
/translation="AELVKGASVKLSCTAGSNFKDITYMHVVKRPEQGLEWIGRID
PANGNTRYDKPFQKGATITADTSSNTAYLQLSLTSEDYAVYCYARGTTVVADYWGQ
T"

CDS
ORIGIN
Query Match      84.3%; Score 43; DB 10; Length 312;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATCCAGGCC 51
    |||||
Db 124 AGGATTGATCTCGGAATGGTAAATATGACCCGAGTTCCAGGCC 174
    |||||

RESULT 29
MMU307470      312 bp DNA linear ROD 07-FEB-2002
LOCUS          Mus musculus partial 17.2.25-D-Jh4 gene for immunoglobulin heavy
DEFINITION     chain, sequence t24b germline.
ACCESSION      AJ307470
VERSION        17.2.25-D-Jh4 gene; immunoglobulin heavy
KEYWORDS       Mus musculus (house mouse)
SOURCE         Mus musculus
ORGANISM       Mus musculus
REFERENCE      1 Toellner K.M., Jenkinson, W.E., Taylor, D.R., Khan, M., Sze, D.M.,
AUTHORS        Sansom, D.M., Vinuesa, C.G. and MacLennan, I.C.
TITLE          Low-level hypermutation in T cell-independent germinal centers
               compared with high mutation rates associated with T cell-dependent
               germinal centers
JOURNAL        J. Exp. Med. 195 (3), 383-389 (2002)
MEDLINE        21686239
PUBMED         11828014
REFERENCE      2 (bases 1 to 312)
AUTHORS        Toellner, K.M.
TITLE          Direct Submission
JOURNAL        Submitted (01-FEB-2001) Toellner K.M., Immunology, University of
               Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED
               KINGDOM
FEATURES       Location/Qualifiers
               1..312
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /cell_type="germinal center"
                /tissue_type="spleen"
                /rearranged
                /note="4 d after NP/CGG immunisation in CGG primed
                recipients of QM B cells"
gene           1..312
               /gene="17.2.25-D-Jh4"
CDS            <1..>312
               /codon_start=1
               /product="immunoglobulin heavy chain"
               /protein_id="CAC29273.1"
               /db_xref="GI:12964070"
               /translation="GAELVKPGASVKLSCTASGPNIKDITYMHVVKRPEQGLEWIGRI
               DPANGNTKYDPKFGKATITADTSSNTAYLQLSLTSEDYAVYCYARYPYVYAMDY
               WG"
gene           1..273
               /gene="17.2.25"

```

```

V_segment      <1..>273
gene           /gene="17.2.25"
J_segment      288..312
               /gene="Jh4"
ORIGIN         288..>312
               /gene="Jh4"
Query Match      84.3%; Score 43; DB 10; Length 312;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATCCAGGCC 51
    |||||
Db 127 AGGATTGATCTCGGAATGGTAAATATGACCCGAGTTCCAGGCC 177
    |||||

RESULT 30
MMU307472      312 bp DNA linear ROD 07-FEB-2002
LOCUS          Mus musculus partial 17.2.25-D-Jh4 gene for immunoglobulin heavy
DEFINITION     chain, sequence t12a.
ACCESSION      AJ307472
VERSION        17.2.25-D-Jh4 gene; immunoglobulin heavy chain.
KEYWORDS       Mus musculus (house mouse)
SOURCE         Mus musculus
ORGANISM       Mus musculus
REFERENCE      1 Toellner K.M., Jenkinson, W.E., Taylor, D.R., Khan, M., Sze, D.M.,
AUTHORS        Sansom, D.M., Vinuesa, C.G. and MacLennan, I.C.
TITLE          Low-level hypermutation in T cell-independent germinal centers
               compared with high mutation rates associated with T cell-dependent
               germinal centers
JOURNAL        J. Exp. Med. 195 (3), 383-389 (2002)
MEDLINE        21686239
PUBMED         11828014
REFERENCE      2 (bases 1 to 312)
AUTHORS        Toellner, K.M.
TITLE          Direct Submission
JOURNAL        Submitted (01-FEB-2001) Toellner K.M., Immunology, University of
               Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED
               KINGDOM
FEATURES       Location/Qualifiers
               1..312
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /cell_type="germinal center"
                /tissue_type="spleen"
                /rearranged
                /note="4 d after NP/CGG immunisation in CGG primed
                recipients of QM B cells"
gene           1..312
               /gene="17.2.25-D-Jh4"
CDS            <1..>312
               /codon_start=1
               /product="immunoglobulin heavy chain"
               /protein_id="CAC29275.1"
               /db_xref="GI:12964074"
               /translation="GAELVKPGASVKLSCTASGPNIKDITYMHVVKRPEQGLEWIGRI
               DPANGNTKYDPKFGKATITADTSSNTAYLQLSLTSEDYAVYCYARYPYVYAMDY
               WG"
gene           1..273
               /gene="17.2.25"
V_segment      <1..>273
gene           /gene="17.2.25"
J_segment      288..312
               /gene="Jh4"
               >312
               /gene="Jh4"

```

## ORIGIN

Query Match 84.3%; Score 43; DB 10; Length 312;  
 Best Local Similarity 90.2%; Pred. No. 3e-05;  
 Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 AAGATTGATCCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51  
 |||||  
 Db 127 AGGATTGATCCTCGGAATGGTAACTAAATATGACCCGAGTTCCAGGCC 177  
 |||||

## RESULT 31

MMU307473

LOCUS

DEFINITION Mus musculus partial 17.2.25-D-Jh4 DNA linear ROD 07-FEB-2002  
 chain, sequence t12b.

ACCESSION

AJ307473

VERSION

AJ307473.1 GI:12964075

KEYWORDS

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Toellner,K.M., Jenkinson,W.E., Taylor,D.R., Khan,M., Sze,D.M.,  
 Sansom,D.M., Vinuesa,C.G. and MacLennan,I.C.

TITLE

Low-level hypermutation in T cell-independent germinal centers  
 compared with high mutation rates associated with T cell-dependent  
 germinal centers

J. Exp. Med. 195 (3), 383-389 (2002)

MEDLINE

21686239

PUBMED

11828014

REFERENCE

2 (bases 1 to 312)

Toellner,K.M.

Direct Submission

TITLE

Submitted (01-FEB-2001) Toellner K.M., Immunology, University of  
 Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED  
 KINGDOM

JOURNAL

FEATURES

source

Location/Qualifiers

1. .312

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/cell\_type="germinal center"

/tissue\_type="spleen"

/rearranged

/notes="4 d after NP/CGG immunisation in CGG primed  
 recipients of QM B cells"

gene

1. .312

/gene="17.2.25-D-Jh4"

CDS

&lt;1. .&gt;312

/gene="17.2.25-D-Jh4"

/codon\_start=1

/product="immunoglobulin heavy chain"

/protein\_id="CAC29276.1"

/db\_xref="GI:12964076"

/translation="GAELVKPGASVKLSCTASGFNIKDTYMWVKORPEGLEWIGRI  
 DPANGNTKYDPKFGKATITADTSSNTAYLQLSSLTSEDYAVYCYRYPYTYMDY  
 WG"

gene

1. .273

/gene="17.2.25"

V\_segment

&lt;1. .273

/gene="17.2.25"

gene

288. .312

/gene="Jh4"

J\_segment

288. .&gt;312

/gene="Jh4"

## ORIGIN

Query Match 84.3%; Score 43; DB 10; Length 312;  
 Best Local Similarity 90.2%; Pred. No. 3e-05;  
 Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy

Db

127

AGGATTGATCCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51

|||||

127

AGGATTGATCCTCGGAATGGTAACTAAATATGACCCGAGTTCCAGGCC 177

|||||

## RESULT 32

MMU307474

LOCUS

DEFINITION Mus musculus partial 17.2.25-D-Jh4 DNA linear ROD 07-FEB-2002  
 chain, sequence t13b.

ACCESSION

AJ307474

VERSION

AJ307474.1 GI:12964077

KEYWORDS

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Toellner,K.M., Jenkinson,W.E., Taylor,D.R., Khan,M., Sze,D.M.,  
 Sansom,D.M., Vinuesa,C.G. and MacLennan,I.C.

TITLE

Low-level hypermutation in T cell-independent germinal centers  
 compared with high mutation rates associated with T cell-dependent  
 germinal centers

J. Exp. Med. 195 (3), 383-389 (2002)

MEDLINE

21686239

PUBMED

11828014

REFERENCE

2 (bases 1 to 312)

Toellner,K.M.

Direct Submission

TITLE

Submitted (01-FEB-2001) Toellner K.M., Immunology, University of  
 Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED  
 KINGDOM

JOURNAL

FEATURES

source

Location/Qualifiers

1. .312

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/cell\_type="germinal center"

/tissue\_type="spleen"

/rearranged

/notes="4 d after NP/CGG immunisation in CGG primed  
 recipients of QM B cells"

gene

1. .312

/gene="17.2.25-D-Jh4"

CDS

&lt;1. .&gt;312

/gene="17.2.25-D-Jh4"

/codon\_start=1

/product="immunoglobulin heavy chain"

/protein\_id="CAC29277.1"

/db\_xref="GI:12964078"

/translation="GAELVKPGASVKLSCTASGFNIKDTYMWVKORPEGLEWIGRI  
 DPANGNTKYDPKFGKATITADTSSNTAYLQLSSLTSEDYAVYCYRYPYTYMDY  
 WG"

gene

1. .273

/gene="17.2.25"

V\_segment

&lt;1. .273

/gene="17.2.25"

gene

288. .312

/gene="Jh4"

J\_segment

288. .&gt;312

/gene="Jh4"

Query Match

Best Local Similarity

Matches

46; Conservative

0; Mismatches

5; Indels

0; Gaps

0;

1 AAGATTGATCCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51

|||||

127 AGGATTGATCCTCGGAATGGTAACTAAATATGACCCGAGTTCCAGGCC 177

|||||

## RESULT 33





```

Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED
KINGDOM
FEATURES
    source      Location/Qualifiers
    1..312      /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /cell_type="germinal center"
                /tissue_type="spleen"
                /rearranged
    note="4 d after NP/CGG immunisation in CGG primed
    recipients of QM B cells"
    1..312      /gene="17.2.25-D-Jh4"
    CDS          /gene="17.2.25-D-Jh4"
                <1..>312
                /codon_start=1
                /product="immunoglobulin heavy chain"
                /protein_id="CAC29282.1"
                /db_xref="GI:12964088"
                /translation="GAEIVKPGASVKLSCTASGFNIDKTYMWVKRPEQGLEWIGRI
                DPANGNTKYPKFGKATITADTSSNTAYLQLSLTSDTAVVYCARYRYPYVAMDY
                WG"
    gene         1..273      /gene="17.2.25"
    V_segment    <1..273      /gene="17.2.25"
                288..312
    gene         /gene="Jh4"
    J_segment    288..>312   /gene="Jh4"
    ORIGIN
    Query Match      84.3%; Score 43; DB 10; Length 312;
    Best Local Similarity 90.2%; Pred. No. 3e-05;
    Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCGGATATTCAGGCC 51
    | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 AGGATTGATCTCGGAATGGTAAATGCTAAATATGACCGGATTCAGGCC 177

RESULT 38
MMU307480
LOCUS      MMU307480      312 bp      DNA      linear      ROD 07-FEB-2002
DEFINITION Mus musculus partial 17.2.25-D-Jh4 gene for immunoglobulin heavy
            chain, sequence t23b.
ACCESSION  AJ307480
VERSION     AJ307480.1 GI:12964089
KEYWORDS    17.2.25-D-Jh4 gene; immunoglobulin heavy chain.
SOURCE      Mus musculus
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS     Toellner,K.M., Jenkinson,W.E., Taylor,D.R., Khan,M., Sze,D.M.,
            Sansom,D.M., Vinuesa,C.G. and MacLennan,I.C.
TITLE       Low-level hypermutation in T cell-independent germinal centers
            compared with high mutation rates associated with T cell-dependent
            germinal centers
JOURNAL     J. Exp. Med. 195 (3), 383-389 (2002)
MEDLINE     21686239
PUBMED      11828014
REFERENCE   2 (bases 1 to 312)
AUTHORS     Toellner,K.M.
TITLE       Direct Submission
JOURNAL     Submitted (01-FEB-2001) Toellner K.M., Immunology, University of
            Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED
            KINGDOM
FEATURES
    source      Location/Qualifiers
    1..312      /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /cell_type="germinal center"
                /tissue_type="spleen"
                /rearranged
    note="4 d after NP/CGG immunisation in CGG primed
    recipients of QM B cells"

Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED
KINGDOM
FEATURES
    source      Location/Qualifiers
    1..312      /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /cell_type="germinal center"
                /tissue_type="spleen"
                /rearranged
    note="4 d after NP/CGG immunisation in CGG primed
    recipients of QM B cells"

```

```

/db_xref="taxon:10090"
/cell_type="germinal center"
/tissue_type="spleen"
/rearranged
/note="4 d after NP/CGG immunisation in CGG primed
recipients of QM B cells"
1..312
/gene="17.2.25-D-Jh4"
<1..>312
/gene="17.2.25-D-Jh4"
/codon_start=1
/product="immunoglobulin heavy chain"
/protein_id="CAC29283.1"
/db_xref="GI:12964090"
/translation="GAELVKGASVKLSCTASGFNIDKTYMHWVKRPEQGLEWIGRI
DPANGNTKYPKFGKATITADTSSNTAYLQLSLTSDTAVVYCARYRYLYSM DY
WG"
1..273
/gene="17.2.25"
<1..273
/gene="17.2.25"
288..312
/gene="Jh4"
288..>312
/gene="Jh4"
ORIGIN
Query Match      84.3%; Score 43; DB 10; Length 312;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCGGATATTCAGGCC 51
    | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 AGGATTGATCTCGGAATGGTAAATGCTAAATATGACCGGATTCAGGCC 177

RESULT 39
MMU307481
LOCUS      MMU307481      312 bp      DNA      linear      ROD 07-FEB-2002
DEFINITION Mus musculus partial 17.2.25-D-Jh4 gene for immunoglobulin heavy
            chain, sequence t24a.
ACCESSION  AJ307481
VERSION     AJ307481.1 GI:12964091
KEYWORDS    17.2.25-D-Jh4 gene; immunoglobulin heavy chain.
SOURCE      Mus musculus
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS     Toellner,K.M., Jenkinson,W.E., Taylor,D.R., Khan,M., Sze,D.M.,
            Sansom,D.M., Vinuesa,C.G. and MacLennan,I.C.
TITLE       Low-level hypermutation in T cell-independent germinal centers
            compared with high mutation rates associated with T cell-dependent
            germinal centers
JOURNAL     J. Exp. Med. 195 (3), 383-389 (2002)
MEDLINE     21686239
PUBMED      11828014
REFERENCE   2 (bases 1 to 312)
AUTHORS     Toellner,K.M.
TITLE       Direct Submission
JOURNAL     Submitted (01-FEB-2001) Toellner K.M., Immunology, University of
            Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED
            KINGDOM
FEATURES
    source      Location/Qualifiers
    1..312      /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /cell_type="germinal center"
                /tissue_type="spleen"
                /rearranged
    note="4 d after NP/CGG immunisation in CGG primed
    recipients of QM B cells"

```

```
gene      1. .312
          /gene="17.2.25-D-Jh4"
CDS       <1. .>312
          /gene="17.2.25-D-Jh4"
          /codon_start=1
          /product="immunoglobulin heavy chain"
          /protein_id="CAC29284.1"
          /db_xref="GI:12964092"
          /translation="GAELVKGASVKLSCTASGFNISKDTYMHVWVQRPQGLEWIGRI
DPANGNTKYDPKFGKATITADTSSNTAYLQLSLTSEDITAVYYCARYRPPYAMDY
WG"
gene      1. .273
          /gene="17.2.25"
V_segment <1. .273
          /gene="17.2.25"
gene      288. .312
          /gene="Jh4"
J_segment 288. .>312
          /gene="Jh4"

ORIGIN
Query Match      84.3%; Score 43; DB 10; Length 312;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
        |||||||
Db      127 AGGATTGATCTCGGAATGGTAAATATGACCCGAGTTCCAGGCC 177

RESULT 40
MMU307482
LOCUS    Mus musculus partial 17.2.25-D-Jh4 gene for immunoglobulin heavy
DEFINITION chain, sequence t32a.
ACCESSION AJ307482
VERSION   AJ307482.1 GI:12964093
KEYWORDS  17.2.25-D-Jh4 gene; immunoglobulin heavy chain.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS   Toellner,K.M., Jenkinson,W.E., Taylor,D.R., Khan,M., Sze,D.M.,
          Sansom,D.M., Vinuesa,C.G. and MacLennan,I.C.
TITLE     Low-level hypermutation in T cell-independent germinal centers
          compared with high mutation rates associated with T cell-dependent
          germinal centers
JOURNAL   J. Exp. Med. 195 (3), 383-389 (2002)
MEDLINE   21686239
PUBMED    11828014
REFERENCE 2 (bases 1 to 312)
AUTHORS   Toellner,K.M.
TITLE     Direct Submission
JOURNAL   Submitted (01-FEB-2001) Toellner K.M., Immunology, University of
          Birmingham, Medical School, Edgbaston, Birmingham, B15 2TT, UNITED
          KINGDOM

FEATURES             Location/Qualifiers
     source           1. .312
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /cell_type="germinal center"
                     /tissue_type="spleen"
                     /rearranged
                     /note="4 d after NP/CGG immunisation in CGG primed
                     recipients of QM B cells"
     gene             1. .312
                     /gene="17.2.25-D-Jh4"
     CDS              <1. .>312
                     /gene="17.2.25-D-Jh4"
                     /codon_start=1
                     /product="immunoglobulin heavy chain"
```

```
/protein_id="CAC29285.1"
/db_xref="GI:12964094"
/translation="GAELVKGASVKLSCTASGFNISKDTYMHVWVQRPQGLEWIGRI
DPANGNTKYDPKFGKATITADTSSNKAYQLSLTSEDITAVYYCARYRPPYAMDY
WG"
gene      1. .273
          /gene="17.2.25"
V_segment <1. .273
          /gene="17.2.25"
gene      288. .312
          /gene="Jh4"
J_segment 288. .>312
          /gene="Jh4"

ORIGIN
Query Match      84.3%; Score 43; DB 10; Length 312;
Best Local Similarity 90.2%; Pred. No. 3e-05;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
        |||||||
Db      127 AGGATTGATCTCGGAATGGTAAATATGACCCGAGTTCCAGGCC 177

Search completed: June 4, 2005, 14:31:36
Job time : 1156.48 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 10:23:32 ; Search time 314.242 Seconds  
(without alignments)  
960.745 Million cell updates/sec

Title: US-10-089-452-25

Perfect score: 51

Sequence: 1 aagattgatctgcgaatg.....atgaccgatattccaggcc 51

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	51	4 AAF88109	AAF88109 H. pylori
2	51	100.0	51	4 AAF88052	AAF88052 H. pylori
3	51	100.0	360	4 AAF88153	AAF88153 H. pylori
4	51	100.0	360	4 AAF88096	AAF88096 H. pylori
5	43	84.3	51	6 ABL53536	ABL53536 Anti-RANK
6	43	84.3	279	9 ACD27478	ACD27478 Mouse DNA
7	43	84.3	357	10 ABZ268619	ABZ268619 Nucleotid
8	43	84.3	399	6 ABL53531	ABL53531 Anti-RANK
9	43	84.3	412	2 AAT73612	AAT73612 CDNA enco
10	43	84.3	414	5 AAH75978	AAH75978 VEGF SCFv
11	43	84.3	423	13 ADS94359	ADS94359 Antibody
12	43	84.3	618	8 ACC44922	ACC44922 TSH recep
13	43	84.3	618	8 ACC44923	ACC44923 TSH recep
14	43	84.3	813	2 AAX56753	AAX56753 Mouse scf
15	43	84.3	1314	2 AAV09256	AAV09256 Nucleotid
16	41.6	81.6	339	2 AAQ71285	AAQ71285 Heavy cha
17	41.4	81.2	132	2 AAQ94536	AAQ94536 Oligonuc
18	41.4	81.2	132	2 AAT38650	AAT38650 Chimaeric
19	41.4	81.2	330	9 ACD27477	ACD27477 Mouse DNA
20	41.4	81.2	351	6 ABZ21829	ABZ21829 Clone the

21	41.4	81.2	351	9 ACC84737	ACC84737 Mouse MAb
22	41.4	81.2	351	12 ADQ31259	ADQ31259 Murine 11
23	41.4	81.2	360	2 AAV02197	AAV02197 cDNA for
24	41.4	81.2	360	6 ABK47434	ABK47434 Mouse cDN
25	41.4	81.2	360	12 ADF70828	ADF70828 NR-LU-13
26	41.4	81.2	409	2 AAQ94485	AAQ94485 Plasmid p
27	41.4	81.2	409	2 AAQ94539	AAQ94539 Human/mur
28	41.4	81.2	409	2 AAT38600	AAT38600 Chimaeric
29	41.4	81.2	409	2 AAT38653	AAT38653 Chimaeric
30	41.4	81.2	436	2 AAQ48001	AAQ48001 V heavy c
31	41.4	81.2	454	2 AAQ99894	AAQ99894 Human VLA
32	41.4	81.2	454	2 AAT74789	AAT74789 Alpha-4 i
33	41.4	81.2	454	13 ADR40431	ADR40431 Reshaped
34	41.4	81.2	454	13 ADR40349	ADR40349 Reshaped
35	41.4	81.2	470	2 AAQ99892	AAQ99892 Mouse VLA
36	41.4	81.2	470	2 AAT74760	AAT74760 Alpha-4 i
37	41.4	81.2	470	13 ADR40418	ADR40418 Mouse 21.
38	41.4	81.2	470	13 ADR40336	ADR40336 Mouse 21.
39	41.4	81.2	822	2 AAQ94548	AAQ94548 Plasmid p
40	41.4	81.2	822	2 AAT38662	AAT38662 Chimaeric
41	41.4	81.2	1344	12 ADQ31269	ADQ31269 Murine 11
42	41.4	81.2	1344	12 ADQ31275	ADQ31275 Humanised
43	41.4	81.2	1344	12 ADQ31273	ADQ31273 Humanised
44	39.8	78.0	51	6 ABK97819	ABK97819 DNA encod
45	39.8	78.0	357	2 AAV43998	AAV43998 Human MAb

#### ALIGNMENTS

##### RESULT 1

AAF88109  
ID AAF88109 standard; DNA; 51 BP.

XX AAF88109;

XX 17-JUL-2001 (first entry)

XX H. pylori catalase derived antibody HP25/6m/1B5 heavy chain CDR2 DNA.

XX Catalase; beta-urease; antibody; antigen; detection; infection; epitope;  
XX acid-resistant microorganism; complementarity determining region; CDR;  
XX feces; heavy chain; light chain; ds.

XX Unidentified.

XX OS

XX WO200127612-A2.

XX 19-APR-2001.

XX 12-OCT-2000; 2000WO-EP010057.

XX 12-OCT-1999; 99EP-00120351.

XX 16-MAR-2000; 2000EP-00105592.

XX 31-MAR-2000; 2000EP-00107028.

XX 10-MAY-2000; 2000EP-00110110.

XX (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

XX Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;

XX WPI; 2001-282086/29.

XX P-PSDB; AAB86082.

XX Detecting infections by acid-resistant microorganisms, particularly for  
XX diagnosing Helicobacter pylori, comprises immunochromatographic detection  
XX of antigen in feces.

XX Claim 26; Page 25; 90pp; German.

XX This invention describes a novel method for detecting infection by an

XX acid-resistant microorganism (A), in a mammal, using immunochromatography. The method is used to diagnose infection by an acid

-resistant microorganism (A), in a mammal, such as Helicobacter, Campylobacter or Mycobacterium, particularly H. pylori (most preferred), H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple, inexpensive and non-invasive, and may indicate the stage of infection. A test strip used in the method may include a filter to eliminate particles present in the sample and only a single receptor provides a reasonably secure diagnosis, with specificity and selectivity improved by detecting several epitopes (of catalase) or different antigens (catalase and beta-urease). The method can be automated. This sequence encodes a complementarity determining region (CDR) from an antibody raised against the H. pylori catalase or beta-urease antigen which is used to illustrate the method of the invention

Query Match 100.0%; Score 51; DB 4; Length 51;  
Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAGATTGATCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 1 AAGATTGATCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51

## RESULT 2

AAF88052  
ID AAF88052 standard; DNA; 51 BP.

AC AAF88052;

DT 17-JUL-2001 (first entry)

DE H. pylori catalase derived antibody HP25/6m/1B5 heavy chain CDR2 DNA.

KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen; infection; acid-resistant microorganism; fecal; antibody; diagnosis; antibacterial; complementarity determining region; db.

OS Unidentified.

PN WO200127613-A2.

PD 19-APR-2001.

PF 12-OCT-2000; 2000WO-EP010058.

PR 12-OCT-1999; 99EP-00120351.

PR 16-MAR-2000; 2000EP-00105592.

PR 31-MAR-2000; 2000EP-00107028.

PR 10-MAY-2000; 2000EP-00110110.

PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

PI Reiter C, Cullmann G, Heppner P, Ringels A, Mueller H, Haindl E;

DR WPI; 2001-282087/29.

DR P-PSDB; AAB86050.

XX Detecting infections by acid-resistant microorganisms, particularly for diagnosing Helicobacter pylori, comprises an immunoassay on a fecal sample.

PS Claim 20; Page 16; 89pp; German.

XX This invention describes a novel method for detecting, in a mammal, infection by an acid-resistant microorganism (A) which comprises reacting a fecal sample with: (i) a receptor (R) such that a complex is formed with an antigen (Ag) of (A); or (ii) two different R so that a three-part complex is formed with Ag, and the formation of a complex detected. R are specific for an Ag which, after passage through the intestines, at least in some mammals, retains a native (or corresponding) structure against which the mammal produces antibodies (when immunized or infected with (A), or its extracts, lysates or derived proteins (or fragments) or

CC synthetic peptides). The products of the invention have antibacterial activity. The method is used to diagnose infection by Helicobacter, Campylobacter or Mycobacterium, particularly H. pylori (most preferred), H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the progress of treatment. Receptors, particularly antibodies, directed against Ag can be used therapeutically for treatment of infections. The method requires only one R to provide a reasonably secure diagnosis (although use of two R improves sensitivity), so is relatively inexpensive and more easily standardized. Also it is direct, non-invasive, suitable for automation and may indicate the stage of an infection. This sequence encodes a complementarity determining region (CDR) from an antibody generated against a Helicobacter pylori antigen (catalase or beta-urease) which is used to illustrate the method of the invention

Query Match 100.0%; Score 51; DB 4; Length 51;  
Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51

Db 1 AAGATTGATCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51

## RESULT 3

AAF88153

ID AAF88153 standard; DNA; 360 BP.

AC AAF88153;

DT 17-JUL-2001 (first entry)

DE H. pylori catalase antibody HP25/6m/1B50 V region heavy chain DNA.

KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope; acid-resistant microorganism; complementarity determining region; CDR; feces; heavy chain; light chain; ds.

OS Unidentified.

PN WO200127612-A2.

PD 19-APR-2001.

PF 12-OCT-2000; 2000WO-EP010057.

PR 12-OCT-1999; 99EP-00120351.

PR 16-MAR-2000; 2000EP-00105592.

PR 31-MAR-2000; 2000EP-00107028.

PR 10-MAY-2000; 2000EP-00110110.

PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.

PI Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;

DR WPI; 2001-282086/29.

DR P-PSDB; AAB86107.

XX Detecting infections by acid-resistant microorganisms, particularly for diagnosing Helicobacter pylori, comprises immunochromatographic detection of antigen in feces.

PS Claim 32; Fig 3; 90pp; German.

XX This invention describes a novel method for detecting infection by an acid-resistant microorganism (A), in a mammal, using immunochromatography. The method is used to diagnose infection by an acid-resistant microorganism (A), in a mammal, such as Helicobacter, Campylobacter or Mycobacterium, particularly H. pylori (most preferred), H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple, inexpensive and non-invasive, and may indicate the stage of infection. A



CC test strip used in the method may include a filter to eliminate particles  
 CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence encodes a  
 CC Helicobacter pylori catalase derived antibody V-region heavy chain  
 CC fragment used to illustrate the method of the invention  
 XX  
 SQ Sequence 360 BP; 87 A; 95 C; 90 G; 88 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 4; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51  
 |||||  
 Db 148 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 198

RESULT 4  
 AAF88096  
 ID AAF88096 standard; DNA; 360 BP.  
 XX AC  
 XX AAF88096;  
 XX  
 DT 17-JUL-2001 (first entry)  
 XX  
 DE H. pylori catalase antibody HP25/6m/1B5 V-region heavy chain DNA.  
 XX  
 KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
 KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
 KW antibacterial; complementarity determining region; V-region; db.  
 XX  
 OS Unidentified.  
 XX  
 XX WO200127613-A2.  
 XX  
 XX 19-APR-2001.  
 XX  
 XX 12-OCT-2000; 2000WO-EP010058.  
 XX  
 PR 12-OCT-1999; 99EP-00120351.  
 PR 16-MAR-2000; 2000EP-00105592.  
 PR 31-MAR-2000; 2000EP-00107028.  
 PR 10-MAY-2000; 2000EP-00110110.  
 XX  
 PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
 XX  
 XX Reiter C, Cullmann G, Heppner P, Ringeis A, Mueller H, Haindl E;  
 XX  
 DR WFI; 2001-282087/29.  
 DR P-PSDB; AAB86069.  
 XX  
 XX Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
 PT sample.  
 XX  
 XX Claim 28; Fig 3; 89pp; German.

CC This invention describes a novel method for detecting, in a mammal,  
 CC infection by an acid-resistant microorganism (A) which comprises reacting  
 CC a fecal sample with: (i) a receptor (R) such that a complex is formed  
 CC with an antigen (Ag) of (A); or (ii) two different R so that a three-part  
 CC complex is formed with Ag, and the formation of a complex detected. R are  
 CC specific for an Ag which, after passage through the intestines, at least  
 CC in some mammals, retains a native (or corresponding) structure against  
 CC which the mammal produces antibodies (when immunized or infected with  
 CC (A), or its extracts, lysates or derived proteins (or fragments) or  
 CC synthetic peptides). The products of the invention have antibacterial  
 CC activity. The method is used to diagnose infection by Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the  
 CC progress of treatment. Receptors, particularly antibodies, directed

CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an  
 CC infection. This sequence encodes a Helicobacter pylori anti-catalase  
 CC derived antibody HP25/6m/1B5 V-region heavy chain fragment which is  
 CC described in the method of the invention  
 XX  
 SQ Sequence 360 BP; 87 A; 95 C; 90 G; 88 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 4; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51  
 |||||  
 Db 148 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 198

RESULT 5  
 ABL53536  
 ID ABL53536 standard; cDNA; 51 BP.  
 XX AC  
 XX ABL53536;  
 XX  
 DT 10-JUN-2002 (first entry)  
 XX  
 DE Anti-RANK ligand mAb 19H22 VH CDR2 DNA.  
 XX  
 KW RANK ligand; RANK-L; monoclonal antibody; mAb; 19H22; antibody;  
 KW rheumatoid arthritis; osteoporosis; bone cancer; metastasis; osteolysis;  
 KW osteoarthritis; psoriasis; diabetes; multiple sclerosis;  
 KW inflammatory bowel disease; cytostatic; antiarthritic; antiinflammatory;  
 KW antidiabetic; osteopathic; antirheumatic; immunosuppressive;  
 KW antipsoriatic; neuroprotective; heavy chain; VH; mouse; gene; ss.  
 XX  
 OS Mus musculus.  
 XX  
 XX Key Location/Qualifiers  
 CDS 1..51  
 FT /\*tag= a  
 FT /product= "VH\_CDR2"  
 FT /partial  
 FT /note= "the CDS does not include a start or a stop codon"  
 XX  
 XX WO200215846-A2.  
 XX  
 XX 28-FEB-2002.  
 XX  
 XX 21-AUG-2001; 2001WO-US026161.  
 XX  
 XX 21-AUG-2000; 2000US-0226524P.  
 XX 07-SEP-2000; 2000US-0230639P.  
 XX  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 XX (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX  
 XX Sweet RW, Tornetta MA, Truneh A, Wattam TA;  
 XX WFI; 2002-242022/29.  
 XX P-PSDB; ABB75622.

XX A mouse monoclonal antibody 19H22 for treating or preventing osteopenic  
 XX disease including rheumatoid arthritis and osteoporosis.  
 XX Disclosure; Page 45; 45pp; English.  
 XX  
 XX The present sequence is that of DNA encoding complementarity determining  
 XX region 2 (CDR2) of the heavy chain variable region (VH, see ABB75617) of  
 XX claimed anti-human RANK ligand (RANK-L) murine monoclonal antibody (mAb)  
 XX 19H22 (isotype IgG2b/kappa). The mAb is produced by hybridoma cells that  
 XX were obtained by immunising CB6 f1 mice with human RANK-L protein, and

CC fusing spleen cells with X63-Ag8.653 myeloma cells. 19H22 mAb is specific  
 CC for human RANK-L, having a binding affinity characterised by a  
 CC dissociation constant of about 10 power -10 M. RANK-L is a member of the  
 CC tumour necrosis factor family of proteins and a key regulator of the  
 CC immune system, bone development and homeostasis. The invention provides  
 CC altered, chimeric and humanised 19H22 antibodies, including Fab or  
 CC Fab'/2 fragments, antibodies comprising the light and heavy chain  
 CC variable regions of 19H22, and antibodies comprising the CDRs of 19H22.  
 CC These are used in methods for the diagnosis, treatment and prevention of  
 CC osteopenic diseases, including rheumatoid arthritis, osteoporosis,  
 CC metastatic and primary bone cancer, wear debris induced osteolysis or  
 CC osteoarthritis, and immune diseases including psoriasis, insulin  
 CC dependent diabetes, inflammatory bowel disease or multiple sclerosis.  
 CC Methods are also provided for the recombinant production of the  
 CC antibodies, using isolated nucleic acids and transfected host (preferably  
 CC mammalian) host cells, and a method of diagnosing conditions associated  
 CC with Th1 T-cell activity or osteoclast development and activation, in  
 CC particular those listed above  
 XX  
 SQ Sequence 51 BP; 16 A; 10 C; 13 G; 12 T; 0 U; 0 Other;

Query Match 84.3%; Score 43; DB 6; Length 51;  
 Best Local Similarity 90.2%; Pred. No. 1.2e-06;  
 Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 AGATTGATCTCGAATGTAATAAATAATGACCCGATATTCAGGCC 51  
 DB 1 AGGATTGATCTCGAATGTAATACTAATAATGACCCGAAGTTCAGGCC 51

RESULT 6  
 ACD27478  
 ID ACD27478 standard; DNA; 279 BP.

AC ACD27478;  
 AC ACD27478;  
 DT 17-SEP-2003 (first entry)  
 DE Mouse DNA encoding anti-LPS monoclonal antibody VH H10.

XX Waap; tyrosine kinase; enzyme; lipopolysaccharide; LPS; Hep1; mouse; ds;  
 KW antibody; heavy chain variable region; VH H10; monoclonal antibody;  
 KW inner core oligosaccharide; phosphorylation; bacterial infection; VH 7-4.  
 XX

OS Mus sp.  
 FH KEY Location/Qualifiers  
 FT CDS 1..279  
 FT /\*tag= a  
 FT /product= "VH H10"  
 FT /partial  
 FT /note= "No start or stop codon shown"

XX US2003017518-A1.  
 XX  
 XX 23-JAN-2003.  
 XX  
 XX 26-JUN-2002; 2002US-00179851.  
 XX  
 XX 26-JUN-2001; 2001US-0300420P.  
 XX

PA (LAMJ/) LAM J.  
 PA (ZHAO/) ZHAO X.  
 XX  
 XX Lam J, Zhao X;  
 XX

XX WPI; 2003-521592/49.  
 DR P-PSDB; ABU62850.  
 DR

XX Assay for modulators of enzymes involved in phosphorylation of the inner  
 PT core oligosaccharide of a lipopolysaccharide, comprises detecting  
 PT phosphorylated lipopolysaccharide in a test sample after adding a  
 PT specific antibody.

XX Disclosure; Fig 14C; 33pp; English.

XX The invention relates to assaying for modulators of an enzyme involved in  
 CC the phosphorylation of the inner core oligosaccharide (Hep1) of  
 CC lipopolysaccharides (LPS), comprises: (a) incubating a test sample  
 CC comprising the enzyme (e.g. the tyrosine kinase Waap), a candidate  
 CC substance and substrates comprising dephosphorylated LPS and a source of  
 CC phosphate; (b) adding at least one antibody that binds to phosphorylated  
 CC LPS while not binding to dephosphorylated LPS; and (c) detecting of the  
 CC phosphorylated LPS in the test sample by measuring the binding of the  
 CC antibody to phosphorylated LPS (by an ELISA assay, enzyme linked  
 CC immunosorbent assay). An increase or decrease in the amount of  
 CC phosphorylated LPS in the test sample in the presence of the candidate  
 CC substance indicates that the candidate substance is a modulator. Also  
 CC included is a kit for performing the method. The method is for assaying  
 CC for modulators, preferably inhibitors, of an enzyme involved in the  
 CC phosphorylation of the inner core oligosaccharide of LPS, wherein a  
 CC decrease in the amount of phosphorylated LPS in the test sample in the  
 CC presence of the candidate substance indicates that the candidate  
 CC substance is an inhibitor. It is useful in conducting a target discovery  
 CC business by providing assay systems using the method of the invention for  
 CC identifying agents by their ability to modulate and enzyme involved in  
 CC the phosphorylation of the inner core oligosaccharide of LPS, conducting  
 CC therapeutic profiling of identified agents for efficacy and toxicity in  
 CC animals, and licensing the rights to a third party for further drug  
 CC development and/or sales or the identified agents or their analogues. The  
 CC method does not require the use of radio-labeled substrates. It is  
 CC amenable to automation. LPS a major virulence factor for infectious  
 CC bacteria such as P. aeruginosa. The present sequence encodes the heavy  
 CC chain variable region from mouse monoclonal antibody H10, the parental  
 CC antibody of 7-4, which recognises the inner polysaccharide core of P.  
 CC aeruginosa LPS  
 XX

SQ Sequence 279 BP; 75 A; 72 C; 73 G; 59 T; 0 U; 0 Other;

Query Match 84.3%; Score 43; DB 9; Length 279;  
 Best Local Similarity 90.2%; Pred. No. 1.6e-06;  
 Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGATTGATCTCGAATGTAATAAATAATGACCCGATATTCAGGCC 51  
 DB 133 AGGATTGATCTCGAATGTAATACTAATAATGACCCGAAGTTCAGGCC 183

RESULT 7  
 ABZ68619  
 ID ABZ68619 standard; DNA; 357 BP.  
 XX  
 AC ABZ68619;  
 XX  
 DT 16-MAY-2003 (first entry)  
 XX  
 DE Nucleotide sequence of K121 antibody heavy chain variable region.  
 XX  
 KW K121 antibody; K121-like antibody; kappa-type myeloma cell;  
 KW kappa-type multiple myeloma; haematopoietic cell transplantation;  
 KW apoptosis; kappa myeloma antigen; gene; ss.  
 XX  
 OS Mus musculus.  
 XX  
 FH KEY Location/Qualifiers  
 FT CDS 1..357  
 FT /\*tag= a  
 FT /partial  
 FT /product= "K121 heavy chain variable region"  
 XX  
 XX WO2003004056-A1.

XX 16-JAN-2003.  
 XX 05-JUL-2002; 2002WO-AU000896.



```

XX PD 20-MAR-1997.
XX PF 11-SEP-1996; 96WO-JP002588.
XX PR 11-SEP-1995; 95JP-00232384.
XX PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX PI Koike M, Furuya A, Nakamura K, Iida A, Anazawa H, Hanai N;
XX PI Takatsu K;
XX DR WPI; 1997-202249/18.
XX DR P-PSDB; AAW21845.
XX PT Antibody against alpha-chain of human interleukin 5 receptor - useful for
XX PT diagnosis and treatment of respiratory allergic diseases, e.g. chronic
XX PT bronchitis.
XX PS Example 2; Page 124-125; 238pp; Japanese.
XX CC The present sequence encodes the heavy chain variable region of the
XX CC murine anti-human interleukin 5 receptor alpha chain (hIL-5R alpha)
XX CC monoclonal antibody (Mab) KM1486. KM1486 is produced by the hybridoma
XX CC FERM BP-5651, which was prepared by immunising Balb/c mice with hIL-5R
XX CC alpha, fusing spleen cells obtained from the mice with mouse myeloma P3-
XX CC U1 cells and screening the resultant hybridomas. The Mab can be used to
XX CC detect or assay for hIL-5R alpha and cells expressing it on their
XX CC surface, especially to diagnose allergic respiratory diseases, e.g.
XX CC chronic bronchitis. It can also be used to treat such diseases. (Updated
XX CC on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 412 BP; 101 A; 107 C; 109 G; 95 T; 0 U; 0 Other;
Query Match 84.3%; Score 43; DB 2; Length 412;
Best Local Similarity 90.2%; Pred. No. 1.7e-06;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy 1 AAGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTCACAGGCC 51
Db 205 AGGATTGATCTCGAATGGTAACTAAATCTGACCCGAGTTCACAGGCC 255
RESULT 10
AAH75978
ID AAH75978 standard; cDNA; 414 BP.
AC AAH75978;
XX 15-NOV-2001 (first entry)
DE VEGF ScFv coding sequence #1.
DE VEGF; ScFv; single chain antibody; antihuman;
KW Vascular Endothelial Growth Factor; tumour; cytostatic; ss.
OS Unidentified.
XX Key Location/Qualifiers
XX CDS 1..414
FT /*tag= a
FT /product= "Human VEGF ScFv"
FT /transl_except= (pos:181..183,aa:Asn)
FT /transl_except= (pos:241..243,aa:Gln)
FT /note= "No start or stop codon given"
XX CN1299833-A.
XX 20-JUN-2001.
XX 30-DEC-1999; 99CN-00117495.
XX 30-DEC-1999; 99CN-00117495.

```

```

XX PA (UYST-) UNIV INST STOMATOMEDICINE HUAXI MEDICINE.
XX PI Wang D, Yang X, Zheng G;
XX DR WPI; 2001-515297/57.
XX DR P-PSDB; AAG66455.
XX PT Human vessel endothelium growth factor resisting single stranded antibody
XX PT and its preparation.
XX PS Claim 1; Page 1 (Claims); 21pp; Chinese.
XX CC The present invention relates to a single chain antibody (ScFv) of
XX CC antihuman Vascular Endothelial Growth Factor (VEGF). The antibody can be
XX CC used for curing tumours. In the present invention, a VEGF synthetic
XX CC peptide was used to establish an antihuman VEGF monoclonal antibody (Mab)
XX CC hybridoma cell strain. The Mab possesses stable bioactivity and obvious
XX CC tumour-inhibiting action. RT-PCR amplification of light and heavy chain
XX CC genes of Mab variable region and DNA fragment was used to implement
XX CC connection to constitute antihuman VEGF ScFv gene and finally obtain the
XX CC high-effective expression of ScFv. The tests showed that the antihuman
XX CC VEGF ScFv is small in molecular weight, possesses strong power for
XX CC penetrating tumour tissue, low immunogenicity and tumour-resisting
XX CC function. The present sequence is a coding sequence for VEGF ScFv, which
XX CC was used in the present invention
XX SQ Sequence 414 BP; 97 A; 105 C; 122 G; 90 T; 0 U; 0 Other;
Query Match 84.3%; Score 43; DB 5; Length 414;
Best Local Similarity 90.2%; Pred. No. 1.7e-06;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy 1 AAGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTCACAGGCC 51
Db 148 AGGATTGATCTCGAATGGTAACTAAATATGACCCGAGTTCACAGGCC 198
RESULT 11
ADS94359
ID ADS94359 standard; DNA; 423 BP.
XX ADS94359;
XX 02-DEC-2004 (first entry)
DE Antibody 16H9 heavy chain variable region (16H9H) DNA SEQ ID NO:58.
XX antibody; antigen-binding antibody fragment;
XX cell-associated CA 125/O772P; monoclonal antibody; cytostatic;
XX immunostimulant; mediator of lysis; tumour; cell proliferative disorder;
XX cancer; cervical cancer; uterine cancer; breast cancer; lung cancer;
XX ovarian cancer; gene; ds.
XX OS Synthetic.
XX Key Location/Qualifiers
XX CDS 1..423
FT /*tag= a
FT /product= "antibody 16H9 heavy chain variable region"
XX WO2004035537-A2.
XX 29-APR-2004.
XX 15-OCT-2003; 2003WO-US032945.
XX 16-OCT-2002; 2002US-0418828P.
XX 10-JUL-2003; 2003US-0485986P.
XX (EURO-) EUROCELTIQUE SA.
XX Albone EF, Soltis DA;

```

XX WPI: 2004-357171/33.  
DR P-PSDB; ADS94356.  
XX  
XX Novel isolated antibody, or antigen-binding antibody fragment binding  
PT with cell-associated CA 125/O772P polypeptide relative to shed CA  
PT 125/O772 polypeptide, useful for ameliorating cervical or ovarian cancer.  
XX  
XX Claim 48; SEQ ID NO 58; 153pp; English.  
XX  
XX The present invention describes an isolated antibody, or an antigen-  
CC binding antibody fragment (I), that preferentially binds cell-associated  
CC CA 125/O772P polypeptide relative to shed CA 125/O772 polypeptide. Also  
CC described: (1) a monoclonal antibody (II) produced by hybridoma 4E7; (2)  
CC a monoclonal antibody that competes with binding of (II); (3) a hybridoma  
CC as deposited in (II); (4) an isolated nucleic acid molecule (III)  
CC comprising a nucleotide sequence that encodes a variable chain region of  
CC (I); (5) a pharmaceutical composition comprising an antibody or an  
CC antigen-binding antibody fragment that preferentially binds cell-  
CC associated CA 125/O772P polypeptide relative to shed CA 125/O772P  
CC polypeptide, and a carrier; (6) a pharmaceutical composition comprising a  
CC monoclonal antibody or an antigen-binding monoclonal antibody fragment  
CC that preferentially binds cell-associated CA 125/O772P polypeptide  
CC relative to shed CA 125/O772P polypeptide, and a carrier; (7) an article  
CC of manufacture (IV) comprising packaging material and a composition  
CC comprising an antibody, or an antigen-binding antibody fragment that  
CC preferentially binds cell-associated CA 125/O772P relative to shed CA  
CC 125/O772P, and a carrier contained within the packaging material, and  
CC composition in a form suitable for administration to a subject; (8) a  
CC fusion polypeptide (V) comprising an antibody, or an antigen-binding  
CC antibody fragment, which preferentially binds cell-associated CA  
CC 125/O772P relative to shed CA 125/O772P operably linked to a heterologous  
CC agent; (9) ameliorating (VI) a symptom of a CA 125/O772P-related disorder  
CC; (10) a monoclonal antibody (VI) chosen from 325.1, 621.1, 633.1, 654.1,  
CC 725.1, 869, 7F10, 8A1, 8AC3, 15C9, 8E3, 8B5, 7G10, 16C7, 7C6, 7H1, 16H9,  
CC 7A11, 4E7, 117.1, 368.1, 446.1, 501.1, and 776.1 or its antigen-binding  
CC antibody fragment; (11) an antibody or antigen binding antibody fragment  
CC that competes with (VI); and (12) a pharmaceutical composition comprising  
CC (VI) and a carrier. (I) has cytostatic activity, and can be used as an  
CC immunostimulant and a mediator of lysis of positive tumour cell. (I) is  
CC useful for ameliorating a symptom of a CA 125/O772P-related disorder which  
CC is a cell proliferative disorder such as cancer, cervical or uterine  
CC cancer, breast or lung cancer or ovarian cancer. (V) is useful  
CC diagnostically for monitoring the development or progression of cancer or  
CC tumour as part of clinical testing procedure. The present sequence  
CC encodes an antibody heavy chain variable region amino acid sequence,  
CC which is used in the exemplification of the present invention.  
XX  
XX Sequence 423 BP; 104 A; 102 C; 118 G; 99 T; 0 U; 0 Other;  
SQ  
Query Match 84.3%; Score 43; DB 13; Length 423;  
Best Local Similarity 90.2%; Pred. No. 1.7e-06;  
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 AAGATTGATCTCGGAATGTAATACTAAATATGACCCGATATTCAGGCC 51  
DB 205 AGGATTGATCTCGGAATGTAATACTAAATATGACCCGATATTCAGGCC 255  
RESULT 12  
ACCA4922  
ID ACCA4922 standard; DNA; 618 BP.  
XX  
XX ACCA4922;  
AC  
XX 05-JUN-2003 (first entry)  
DT  
XX TSH receptor antibody 3B3 heavy chain nucleotide sequence.  
DE  
XX Thyrotropin receptor; TSH receptor; epitope region; antibody;  
KW autoantibody; autoimmune disease; thyroid tissue; cytostatic;  
KW gene therapy; immune reaction; thyroid cancer; gene; ds.  
XX

OS Synthetic.  
XX WO2003018632-A2.  
XX  
XX 06-MAR-2003.  
XX  
XX 21-AUG-2002; 2002WO-GB003831.  
XX  
XX 23-AUG-2001; 2001GB-00020649.  
PR 01-JUL-2002; 2002GB-00015212.  
XX (RSRR-) RSR LTD.  
XX Smith BR, Furmaniak J, Sanders JF;  
PI P-PSDB; ABP96767.  
XX WPI: 2003-290051/28.  
DR P-PSDB; ABP96767.  
XX  
XX New polypeptide sequence having part or all of the primary structural  
PT conformation of one or more TSH receptor epitopes, useful for treating an  
PT autoimmune disease associated with an immune reaction to a TSH receptor,  
PT e.g. thyroid cancer.  
XX  
XX Claim 71; Fig 53; 196pp; English.  
XX  
XX The present invention describes a polypeptide sequence comprising part or  
CC all of the primary structural conformation of one or more thyrotropin  
CC (TSH) receptor epitopes with which autoantibodies and/or lymphocytes  
CC produced in response to a TSH receptor interact. Also described: (1) one  
CC or more receptor TSH epitopes with which the autoantibodies and/or  
CC lymphocytes interact, as with the polypeptide sequence described above;  
CC (2) a method of screening for autoantibodies or lymphocytes produced in  
CC response to a TSH receptor in a sample of body fluid obtained from a  
CC subject suspected of suffering from, susceptible to, having or recovering  
CC from autoimmune disease associated with an immune reaction to a TSH  
CC receptor; (3) a binding partner for a TSH receptor, which is capable of  
CC binding to a TSH receptor to stimulate the TSH receptor, where the  
CC binding partner does not comprise TSH or naturally produced antibodies to  
CC the TSH receptor; and (4) a combination comprising the binding partner  
CC and one or more further agents capable of stimulating thyroid tissue,  
CC and/or tissue containing a TSH receptor, for simultaneous, separate or  
CC sequential use in stimulating thyroid tissue, and/or tissue containing a  
CC TSH receptor. A TSH receptor has cytostatic activity and can be used in  
CC gene therapy. The polypeptide, compositions and methods from the present  
CC invention can be used for treating an autoimmune disease associated with  
CC an immune reaction to a TSH receptor. The specific binding partner is  
CC useful for the manufacture of a medicament for stimulating thyroid tissue  
CC or tissue containing a TSH receptor, and for treating thyroid cancer.  
CC ACC44874 to ACC44905 and ABP96719 to ABP96750 represent TSH receptor  
CC sequences, and ACC44910 to ACC44933 encode the TSH receptor antibody VH  
CC and VL domains given in ABP96751 to ABP96778, which are used in the  
CC exemplification of the present invention  
XX  
XX Sequence 618 BP; 144 A; 194 C; 154 G; 126 T; 0 U; 0 Other;  
SQ  
Query Match 84.3%; Score 43; DB 8; Length 618;  
Best Local Similarity 90.2%; Pred. No. 1.8e-06;  
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 AAGATTGATCTCGGAATGTAATACTAAATATGACCCGATATTCAGGCC 51  
DB 148 AGGATTGATCTCGGAATGTAATACTAAATATGACCCGATATTCAGGCC 198  
RESULT 13  
ACCA4923  
ID ACCA4923 standard; DNA; 618 BP.  
XX  
XX ACCA4923;  
AC  
XX 05-JUN-2003 (first entry)  
DT  
XX TSH receptor antibody 3B3 heavy chain nucleotide sequence.  
DE

XX Thyrotropin receptor; TSH receptor; epitope region; antibody;  
 KW autoantibody; autoimmune disease; thyroid tissue; cytostatic;  
 KW gene therapy; immune reaction; thyroid cancer; gene; ds.  
 XX Synthetic.

XX WO2003018632-A2.

XX 06-MAR-2003.

XX 21-AUG-2002; 2002WO-GB003831.

XX 23-AUG-2001; 2001GB-00020649.

XX 01-JUL-2002; 2002GB-00015212.

XX (RSRR-) RSR LTD.

XX Smith BR, Furmaniak J, Sanders JF;  
 XX P-PSDB; ABP96768.

XX WPI; 2003-290051/28.

XX P-PSDB; ABP96768.

XX New polypeptide sequence having part or all of the primary structural  
 PT conformation of one or more TSH receptor epitopes, useful for treating an  
 PT autoimmune disease associated with an immune reaction to a TSH receptor,  
 PT e.g. thyroid cancer.

XX Claim 71; Fig 54; 196pp; English.

XX The present invention describes a polypeptide sequence comprising part or  
 CC all of the primary structural conformation of one or more thyrotropin  
 CC (TSH) receptor epitopes with which autoantibodies and/or lymphocytes  
 CC produced in response to a TSH receptor interact. Also described: (1) one  
 CC or more receptor TSH epitopes with which the autoantibodies and/or  
 CC lymphocytes interact, as with the polypeptide sequence described above;  
 CC (2) a method of screening for autoantibodies or lymphocytes produced in  
 CC response to a TSH receptor in a sample of body fluid obtained from a  
 CC subject suspected of suffering from, susceptible to, having or recovering  
 CC from autoimmune disease associated with an immune reaction to a TSH  
 CC receptor; (3) a binding partner for a TSH receptor, which is capable of  
 CC binding to a TSH receptor to stimulate the TSH receptor, where the  
 CC binding partner does not comprise TSH or naturally produced antibodies to  
 CC the TSH receptor; and (4) a combination comprising the binding partner  
 CC and one or more further agents capable of stimulating thyroid tissue,  
 CC and/or tissue containing a TSH receptor, for simultaneous, separate or  
 CC sequential use in stimulating thyroid tissue, and/or tissue containing a  
 CC TSH receptor. A TSH receptor has cytostatic activity and can be used in  
 CC gene therapy. The polypeptide, compositions and methods from the present  
 CC invention can be used for treating an autoimmune disease associated with  
 CC an immune reaction to a TSH receptor. The specific binding partner is  
 CC useful for the manufacture of a medicament for stimulating thyroid tissue  
 CC or tissue containing a TSH receptor, and for treating thyroid cancer.  
 CC AC44874 to ACC44905 and ABP96719 to ABP96750 represent TSH receptor  
 CC sequences, and ACC44910 to ACC44933 encode the TSH receptor antibody VH  
 CC and VL domains given in ABP96751 to ABP96778, which are used in the  
 CC exemplification of the present invention

SQ Sequence 618 BP; 144 A; 194 C; 154 G; 126 T; 0 U; 0 Other;

Query Match 84.3%; Score 43; DB 8; Length 618;  
 Best Local Similarity 90.2%; Pred. No. 1.8e-06;  
 Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGAGTTGATCTCGGAATGTAATACTAAATATGACCCGATATTCAGGCC 51  
 DB 148 AGGATTGATCTCGGAATGTAATACTAAATATGACCCGATATTCAGGCC 198

RESULT 14

AA56753

ID AA56753 standard; DNA; 813 BP.

XX

AC AA56753;

XX 14-JUL-1999 (first entry)

XX Mouse scFV fragment 1.

XX SCFV; murine; VH domain; variable region; heavy chain; immunisation;  
 KW receptor-ligand complex; immunoselection; cell proliferation; therapy;  
 KW anti-inflammatory; anticancer; antileukemic; diagnosis; inflammation;  
 KW cancer; carcinoma; breast; stomach; prostate; lung; colon; pancreas;  
 KW Kaposi's sarcoma; leukemia; ss.

XX Mus sp.

XX WO9919361-A1.

XX 22-APR-1999.

XX 08-OCT-1998; 98WO-EP006386.

XX 09-OCT-1997; 97DB-01044531.

XX (BOSS/) BOSSLET K.

XX Bosslet K, Petrus H;

XX WPI; 1999-287955/24.

XX Binding molecules specific for receptor-ligand complex.

XX Claim 17; Page 27; 38pp; German.

XX This invention describes a novel method for binding molecules against a  
 CC receptor-ligand complex produced by immunisation, or immunoselection,  
 CC using the receptor-ligand complex, in which the components are attached  
 CC by at least one covalent bond. The specific binding interaction with the  
 CC activated receptor is involved in cell proliferation. The method has  
 CC antiinflammatory, anticancer and antileukemic applications. The binding  
 CC molecules used in the method of the invention can be optionally labeled  
 CC or attached to a toxin or other active agent and are used in therapy and  
 CC diagnosis of inflammation, solid cancers (e.g. carcinoma of breast,  
 CC stomach, prostate, lung, colon, and pancreas, or Kaposi's sarcoma) and  
 CC leukemia. The binding molecules recognize an epitope present on the  
 CC receptor-ligand complex but not on its separate components, i.e. they are  
 CC specific for activated receptors and thus for proliferative tissue. This  
 CC sequence represents a fragment of the mouse scFv variable region heavy  
 CC chain which is used in the method of the invention

SQ Sequence 813 BP; 202 A; 214 C; 217 G; 180 T; 0 U; 0 Other;

Query Match 84.3%; Score 43; DB 2; Length 813;  
 Best Local Similarity 90.2%; Pred. No. 1.8e-06;  
 Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGAGTTGATCTCGGAATGTAATACTAAATATGACCCGATATTCAGGCC 51  
 DB 153 AGGATTGATCTCGGAATGTAATACTAAATATGACCCGATATTCAGGCC 203

RESULT 15

AAV09256

ID AAV09256 standard; cDNA; 1314 BP.

XX AAV09256;

XX 17-OCT-2003 (revised)

XX 07-JUL-1998 (first entry)

XX Nucleotide sequence encoding the Mgr6-clavin immunotoxin.

XX Recombinant ribosomal inhibitor protein; RIP; clavin; inhibition; ss;

XX protein synthesis; conjugate; Mgr6-clavin; anticancer;

XX antiviral agent fusion protein; chimeric.



```

XX 17-JAN-1996 (first entry)
XX DT
XX Oligonucleotide RVH 3.
XX DE
XX Human; murine; HEP-RVH-M21-g(gamma)1; antibody; diagnosis;
XX KW oligonucleotide RVH 3; medulloblastoma; brain tumour; treatment; ss.
XX KW
XX OS Homo sapiens.
XX PN WO9514041-A1.
XX PD 26-MAY-1995.
XX PF 19-OCT-1994; 94WO-JP001763.
XX PR 19-NOV-1993; 93JP-00291078.
XX PA (CHUS ) CHUGAI SEIYAKU KK.
XX PI Ohtomo T, Sato K, Teuchiya M;
XX DR WPI; 1995-200347/26.
XX PT Reconstituted antibody against human medullo:blastoma cells - contains
XX PT high proportion of human antibody origin and has low antigenicity.
XX PS
XX PS Example 5; Page 92; 120pp; Japanese.
XX CC AAQ94533-Q94536 are oligonucleotides used in the construction of the
XX CC human/murine chimeric antibody HEP-RVH-M21-g(gamma)1. The antibody is
XX CC reactive with human medulloblastoma (a brain tumour) cells. The chimeric
XX CC antibody can be used in the diagnosis and treatment of this disease
XX CC
XX SQ Sequence 132 BP; 25 A; 32 C; 35 G; 40 T; 0 U; 0 Other;

Query Match 81.2%; Score 41.4; DB 2; Length 132;
Best Local Similarity 88.2%; Pred. No. 5.7e-06;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGGAATGTAATAAATAATATGACCCGATATTCAGGCGC 51
Db 132 AAGATTGATCCTCGGAATGTAATAAATAATATGACCCGGAAGTTCAGGCGC 82

RESULT 18
AAT38650/c
ID AAT38650 standard; DNA; 132 BP.
XX AC AAT38650;
XX DT 09-DEC-1996 (first entry)
XX DE Chimaeric human/murine MAb ONS-M21 fragment RVH3.
XX KW Murine; human; myeloblastoma; chimaera; monoclonal antibody; chimera;
XX KW single stranded Fv region; low human antigenicity; diagnosis; treatment;
XX KW cerebral tumour; reshaped; ss.
XX OS Synthetic.
XX PN JP08169900-A.
XX PD 02-JUL-1996.
XX PF 18-NOV-1994; 94JP-00285057.
XX PR 19-NOV-1993; 93JP-00291078.
XX PR 18-OCT-1994; 94JP-00252166.
XX PA (CHUS ) CHUGAI PHARM CO LTD.
XX PI WPI; 1996-358509/36.

```

```

XX Reshaped anti-human myeloblastoma cell human antibody - has low human
XX PT antigenicity, and is therefore useful for diagnosis and treatment of
XX PT cerebral tumours, e.g. myeloblastoma.
XX PS
XX PS Example 5; Page 38; 45pp; Japanese.
XX CC The present sequence is a fragment of the chimaeric human/murine
XX CC monoclonal antibody (MAb) ONS-M21 DNA. The MAb was prepd. by combining
XX CC light and heavy variable region DNA, from a murine anti-human
XX CC myeloblastoma cell MAb, with human light and heavy constant region
XX CC sequences, respectively to produce chimaeric human/murine light and heavy
XX CC chain DNA mols. . A recombinant vector for the expression of the heavy and
XX CC light chain DNA mols. was prepd., and used to transform a host cell. The
XX CC host cell was then cultured, and the expression prods. of the heavy and
XX CC light chain DNA mols. sepd. and connected with a peptide linker to
XX CC produce a single stranded Fv region. The reshaped Fv region has low human
XX CC antigenicity, and is therefore expected to be useful as an agent for the
XX CC diagnosis and treatment of cerebral tumours, e.g. myeloblastoma
XX SQ Sequence 132 BP; 25 A; 32 C; 35 G; 40 T; 0 U; 0 Other;

Query Match 81.2%; Score 41.4; DB 2; Length 132;
Best Local Similarity 88.2%; Pred. No. 5.7e-06;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGGAATGTAATAAATAATATGACCCGATATTCAGGCGC 51
Db 132 AAGATTGATCCTCGGAATGTAATAAATAATATGACCCGGAAGTTCAGGCGC 82

RESULT 19
ACD27477
ID ACD27477 standard; DNA; 330 BP.
XX AC ACD27477;
XX DT 17-SEP-2003 (first entry)
XX DE Mouse DNA encoding anti-LPS monoclonal antibody VH 7-4.
XX KW WaaP; tyrosine kinase; enzyme; lipopolysaccharide; LPS; HcpI; mouse; ds;
XX KW antibody; heavy chain variable region; VH 7-4; monoclonal antibody;
XX KW inner core oligosaccharide; phosphorylation; bacterial infection.
XX OS Mus sp.
XX PH Key Location/Qualifiers
XX FT CDS 1..330
XX FT /*tag= a
XX FT /product= "VH 7-4"
XX FT /partial
XX FT /note= "No start or stop codon shown"
XX PN US2003017518-A1.
XX PD 23-JAN-2003.
XX PF 26-JUN-2002; 2002US-00179851.
XX PR 26-JUN-2001; 2001US-0300420P.
XX PA (LAMJ/) LAM J.
XX PA (ZHAO/) ZHAO X.
XX PI Lam J, Zhao X;
XX DR WPI; 2003-521592/49.
XX DR P-PSDB; ABU62845.
XX PT Assay for modulators of enzymes involved in phosphorylation of the inner
XX PT core oligosaccharide of a lipopolysaccharide, comprises detecting a
XX PT phosphorylated lipopolysaccharide in a test sample after adding a

```



specific antibody.

Disclosure; Fig 14A; 33pp; English.

The invention relates to assaying for modulators of an enzyme involved in the phosphorylation of the inner core oligosaccharide (HepI) of lipopolysaccharides (LPS), comprises: (a) incubating a test sample comprising the enzyme (e.g. the tyrosine kinase Waap), a candidate substance and substrates comprising dephosphorylated LPS and a source of phosphate; (b) adding at least one antibody that binds to phosphorylated LPS while not binding to dephosphorylated LPS; and (c) detecting phosphorylated LPS in the test sample by measuring the binding of the antibody to phosphorylated LPS (by an ELISA assay, enzyme linked immunosorbent assay). An increase or decrease in the amount of phosphorylated LPS in the test sample in the presence of the candidate substance indicates that the candidate substance is a modulator. Also included is a kit for performing the method. The method is for assaying for modulators, preferably inhibitors, of an enzyme involved in the phosphorylation of the inner core oligosaccharide of LPS, wherein a decrease in the amount of phosphorylated LPS in the test sample in the presence of the candidate substance indicates that the candidate substance is an inhibitor. It is useful in conducting a target discovery business by providing assay systems using the method of the invention for identifying agents by their ability to modulate and enzyme involved in the phosphorylation of the inner core oligosaccharide of LPS, conducting therapeutic profiling of identified agents for efficacy and toxicity in animals, and licensing the rights to a third party for further drug development and/or sales or the identified agents or their analogues. The method does not require the use of radio-labeled substrates. It is amenable to automation. LPS a major virulence factor for infectious bacteria such as *P. aeruginosa*. The present sequence encodes the heavy chain variable region from mouse monoclonal antibody 7-4, which recognises the inner polysaccharide core of *P. aeruginosa* LPS

Sequence 330 BP; 86 A; 83 C; 89 G; 72 T; 0 U; 0 Other;

Query Match 81.2%; Score 41.4; DB 9; Length 330;  
Best Local Similarity 88.2%; Pred. No. 6.6e-06;  
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 AAGATTGATCTCGGAATGGTAAACTAAATATATGACCCGATATTCAGGCC 51  
133 AGGATTGATCTCGGAATGGTAACTACTAAATATGACCCGAGTTCCAGGCC 183

RESULT 20  
AB221829  
ID AB221829 standard; DNA; 351 BP.  
XX  
AC AB221829;  
XX  
DT 28-MAR-2003 (first entry)  
XX  
DE Clone therap.152C3 VH sequence.  
XX  
KW Mitochondrial membrane permeabilisation; mitochondrion; PTPC; chimeric;  
KW permeability transition pore complex; virucide; neuroprotective;  
KW vasotrophic; cytosatic; infection; cell death regulation; apoptosis;  
KW mitochondrial permeability transition pore complex modulator; cancer;  
KW apoptogenic; ischaemia; neurodegenerative disease; fulminant hepatitis;  
KW gene; ds.  
XX  
OS Mus sp.  
OS Synthetic.  
XX  
XX WO200261105-A2.  
XX  
XX 08-AUG-2002.  
XX  
XX 01-FEB-2002; 2002WO-EP001633.  
XX  
XX 02-FEB-2001; 2001US-0265594P.  
XX

(INSP ) INST PASTEUR.  
(CNRS ) CENT NAT RECH SCI.  
Edelman L, Jacotot E, Briand J;  
WPI; 2002-619260/66.  
New chimeric bifunctional molecules that target specific cells and regulate the apoptosis function of the permeability transition pore complex of the mitochondria, useful for treating or preventing e.g. cancer or ischemia.  
Example 2; Fig 5; 76pp; English.  
The present invention describes a chimeric bifunctional molecule (I) comprising at least a first functional molecule covalently linked to a second functional molecule, which is able to modulate the activity of the permeability transition pore complex (PTPC) of the mitochondria. (I) has the function of specifically targeting and entering a tissue cell population. The second functional molecule has the function of specifically targeting, and inducing or preventing the death of the cells by apoptosis by regulating the opening or the closing of the PTPC of the mitochondria or its fragment. (I) has virucide, neuroprotective, vasotrophic and cytostatic activities, and can be used as a mitochondrial permeability transition pore complex (PTPC) modulator. (I) is also treating or preventing a pathological infection or disease. (I) is also useful for regulating cell death regulatory molecules, specifically the apoptogenic function of the PTPC, for treating e.g. cancer, ischaemia, neurodegenerative diseases, fulminant hepatitis or viral infections. The present sequence represents an ScFv VH sequence which is used in an example from the present invention

Sequence 351 BP; 84 A; 94 C; 91 G; 82 T; 0 U; 0 Other;

Query Match 81.2%; Score 41.4; DB 6; Length 351;  
Best Local Similarity 88.2%; Pred. No. 6.6e-06;  
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 AAGATTGATCTCGGAATGGTAAACTAAATATATGACCCGATATTCAGGCC 51  
148 AGGATTGATCTCGGAATGGTAAACTATATATATGACCCGAGTTCCAGGCC 198

RESULT 21  
ACC84737  
ID ACC84737 standard; DNA; 351 BP.  
XX  
AC ACC84737;  
XX  
DT 12-SEP-2003 (first entry)  
XX  
DE Mouse MAb 11K2 heavy chain variable region nucleotide sequence.  
XX  
KW Antibody; beta-chemokine; monocyte chemotactic protein; MCP-1; MCP-2;  
KW MCP-3; antiinflammatory; nephrotrophic; dermatological; hepatotrophic;  
KW neuroprotective; antiatherosclerotic; antirheumatic; antiarthritic;  
KW cytosatic; anticonvulsant; antidiabetic; ophthalmological; vasotrophic;  
KW immunomodulator; cerebroprotective; antiulcer; antiparkinsonian; MAB;  
KW monoclonal antibody; 11K2; ds.  
XX  
OS Mus musculus.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..351  
XX /\*tag= a  
XX  
XX WO2003048083-A2.  
XX  
XX 12-JUN-2003.  
XX  
XX 27-NOV-2002; 2002WO-US038229.  
XX  
XX 30-NOV-2001; 2001US-0343391P.  
XX

```

PR 24-MAY-2002; 2002US-0383277P.
XX 01-AUG-2002; 2002US-0400469P.
XX (BIOJ ) BIOGEN INC.
XX
XX De Fougereolles AR, Kotelianski VE, Garber E, Reid C;
XX
XX WPI; 2003-532819/50.
DR P-PSDB; ABR61872.
XX
XX Novel antibody or its antigen binding fragment that specifically binds to
XX monocytic chemotactic protein-2 and other beta-chemokine, useful for
XX detecting presence of beta-chemokine in sample, and blocking chemotaxis.
XX
XX Claim 49; Page 68-69; 100pp; English.
XX
XX The invention relates to an antibody or its antigen binding fragment (AF)
XX that specifically binds to a plurality of beta-chemokines, particularly
XX monocytic chemotactic proteins MCP-1, MCP-2 and MCP-3. The antibodies are
XX useful in an immunoassay method for detecting the presence of a beta-
XX chemokine in a sample, and for blocking chemotaxis. The antibodies or
XX their Afs are useful for treating a subject suffering from a disorder
XX chosen from glomerulonephritis, scleroderma, multiple sclerosis, lupus
XX nephritis, cirrhosis, atherosclerosis, inflammatory bowel disease or
XX rheumatoid arthritis. The antibodies are also useful for treating
XX oncogenic disease or cancer, Crohn's disease, diabetic nephropathy,
XX cachexia, stroke, ulcerative colitis, coronary restenosis, Huntington's
XX disease and Parkinson's disease. The present sequence represents a murine
XX monoclonal antibody (Wab) 11K2 heavy chain variable region nucleotide
XX sequence
XX
XX Sequence 351 BP; 86 A; 94 C; 93 G; 78 T; 0 U; 0 Other;
XX
XX Query Match 81.2%; Score 41.4; DB 9; Length 351;
XX Best Local Similarity 88.2%; Pred. No. 6.6e-06;
XX Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX 1 AAGATTGATCCTCGAATGGTAAACCTAAATATGACCCGATATTCAGGCC 51
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 148 AGGATTGATCCTCGAATGGTAAATCTAAATTTGACCCGAGTTCAGGCC 198
XX
XX RESULT 23
XX ADQ31259
XX ID ADQ31259 standard; DNA; 351 BP.
XX AC ADQ31259;
XX
XX 09-SEP-2004 (first entry)
XX
XX Murine 11K2 heavy chain variable domain antibody DNA SeqID 25.
XX
XX ds; murine; 11K2; monocyte chemotactic protein; gene;
XX beta-chemokine family; MCP-3; glomerulonephritis; scleroderma; cirrhosis;
XX multiple sclerosis; lupus nephritis; atherosclerosis;
XX inflammatory bowel disease; rheumatoid arthritis; inflammatory disease;
XX fibrotic disorder; cancer; immunopathological disorder;
XX antiarteriosclerotic; cancer; immunopathological disorder;
XX cytotatic; dermatological; antiarthritic; antiinflammatory; antirheumatic;
XX neuroprotective; mouse; MCP.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 1..351
XX /*tag= a
XX /partial
XX /product= "11K1 heavy chain variable antibody protein"
XX /note= "Start and stop codons are absent"
XX
XX WO2004050836-A2.
XX
XX 17-JUN-2004.
XX
XX
XX 25-NOV-2003; 2003WO-US037834.
XX
XX 27-NOV-2002; 2002US-0430007P.
XX
XX (BIOG-) BIOGEN IDEC MA INC.
XX
XX De Fougereolles AR, Kotelianski VE, Garber E, Reid C, Saldanha JW;
XX Van Vlijmen H;
XX
XX WPI; 2004-461110/43.
DR P-PSDB; ADQ31261.
XX
XX New antibodies against monocyte chemotactic proteins (MCP), useful for
XX treating or preventing disorders associated with detrimental MCP
XX activity, e.g. glomerulonephritis, scleroderma, multiple sclerosis, or
XX atherosclerosis.
XX
XX Example 10; SEQ ID NO 25; 200pp; English.
XX
XX This invention relates to an antibody for treating or preventing
XX disorders associated with detrimental monocyte chemotactic protein (MCP)
XX activity. Specifically, it refers to humanised antibodies that bind to
XX members of the beta-chemokine family (of which MCP-1, MCP-2 and MCP-3
XX belong) and in particular antibodies that have been modelled on, and
XX modified from, the variable complementarity determining regions (CDRs) of
XX the murine 11K2 and 1A1 immunoglobulin sequences. The present invention
XX describes using these antibodies to treat or prevent diseases and
XX disorders including glomerulonephritis, scleroderma, cirrhosis, multiple
XX sclerosis, lupus nephritis, atherosclerosis, inflammatory bowel diseases,
XX rheumatoid arthritis, inflammatory diseases, fibrotic disorders, cancer
XX and immunopathological disorders. Accordingly, they can be used in the
XX development of pharmaceutical compositions that exhibit
XX antiarteriosclerotic, antiarthritic, antiinflammatory, antirheumatic,
XX cytotatic, dermatological, hepatotropic, immunomodulator, nephrotropic
XX and neuroprotective activities. This polynucleotide sequence is the
XX murine 11K1 heavy chain antibody DNA sequence of the invention.
XX
XX Sequence 351 BP; 86 A; 94 C; 93 G; 78 T; 0 U; 0 Other;
XX
XX Query Match 81.2%; Score 41.4; DB 12; Length 351;
XX Best Local Similarity 88.2%; Pred. No. 6.6e-06;
XX Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX 1 AAGATTGATCCTCGAATGGTAAACCTAAATATGACCCGATATTCAGGCC 51
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 148 AGGATTGATCCTCGAATGGTAAATCTAAATTTGACCCGAGTTCAGGCC 198
XX
XX RESULT 23
XX AAV02197
XX ID AAV02197 standard; cDNA; 360 BP.
XX
XX AAV02197;
XX
XX 25-MAR-2003 (revised)
XX 05-JUN-1998 (first entry)
XX
XX cDNA for heavy chain variable region of NR-LU-13 antibody.
XX
XX Heavy chain; variable region; murine; mouse; human; cancer antigen;
XX antibody; humanised; NR-LU-13; NRX451; cancer; diagnosis; treatment; ss.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX mat_peptide 1..360
XX /*tag= a
XX
XX WO9746589-A2.
XX
XX 11-DEC-1997.
XX
XX

```

```

PP 06-JUN-1997; 97WO-US010074.
XX
PR 07-JUN-1996; 96US-00660362.
XX
XX (NEOR-) NEORX CORP.
XX
PI Graves SS, Reno JM, Mallett RW, Hylarides MD, Searle SMJ;
PI Henry AH, Pedersen JT, Rees AR;
XX
DR WPI; 1998-042124/04.
DR P-PSDB; RAW44121.
XX
XX Humanised antibody binds same human cancer antigen as antibody NR-LU-13 -
PT useful for pre-targeting methods, conventional antibody therapy and
FT immunodiagnosis.
XX
XX Example 1; Fig 2; 100pp; English.
XX
CC The present sequence encodes the heavy chain variable region of the
CC murine anti-human cancer antigen antibody (Ab) NR-LU-13. A novel
CC humanised Ab (hAb) binds the antigen bound by NR-LU-13. The hAb,
CC specifically NR451, or its conjugates can be used for the manufacture of
CC a diagnostic or medicament for cancer diagnosis or treatment. The hAb has
CC reduced immunogenicity and toxicity in humans, but retains the ability to
CC bind the NR-LU-13 antigen. (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 360 BP; 90 A; 92 C; 99 G; 79 T; 0 U; 0 Other;
Query Match 81.2%; Score 41.4; DB 2; Length 360;
Best Local Similarity 88.2%; Pred. No. 6.7e-06;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 AAGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 148 AGGATTGATCTCGAATGGTAACTAAATGTAATGACCCGAAGTTCAGGCC 198

RESULT 24
ABK47434
ID ABK47434 standard; cDNA; 360 BP.
XX
AC ABK47434;
XX
DT 18-JUN-2002 (first entry)
XX
DE Mouse cDNA encoding antibody NR-LU-13 heavy chain variable region.
XX
KW Mouse; ss; gene; NR-LU-13; humanised antibody; NX451; cancer; cytostatic;
KW heavy chain variable region.
XX
OS Mus sp.
XX
XX Key Location/Qualifiers
XX CDS 1..360
XX FT /*tag= a
XX FT /product= "NR-LU-13 heavy chain variable region"
XX FT /partial
XX FT /note= "No start or stop codon shown"
XX
XX US6358710-B1.
XX
XX 19-MAR-2002.
XX
XX 09-JUN-1997; 97US-00871488.
XX
XX 07-JUN-1996; 96US-00660362.
XX
XX (NEOR-) NEORX CORP.
XX
XX Graves SS, Reno JM, Mallett RW, Hylarides MD, Searle SMJ;
XX PI Henry AH, Pedersen JT, Rees AR;
XX
XX WPI; 2002-314754/35.

```

```

DR P-PSDB; AAU79031.
XX
XX Humanized antibody NRX-451, which is derived from ATCC CRL-12360, but
PT which has reduced immunogenicity or toxicity, useful for diagnosing and
PT treating cancers.
XX
XX Example 1; Fig 2; 53pp; English.
XX
CC The invention relates to the humanised antibody NRX 451 (or an antigen-
CC binding fragment), in which the murine parent antibody is ATCC CRL-12360
CC (also known as NU-LU-13 which targets a 40 kilodalton glycoprotein
CC expressed by many carcinomas) and which comprises the light chain
CC variable region appearing as AAU79033 and the heavy chain variable region
CC appearing as AAU79034. The antibody may be administered to treat cancers
CC and use to detect the presence of the antigen in clinical samples. The
CC antibody is a humanised antibody derived from NR-LU-13 antibodies, which
CC exhibits reduced immunogenicity or toxicity in humans but retains the
CC ability to bind the NR-LU-13 antigen. The antibody either does not
CC possess N-linked glycosylation or its N-linked glycosylation has been
CC modified post expression to reduce immunogenicity or toxicity. The
CC present sequence encodes the heavy chain variable region of mouse
CC monoclonal antibody NR-LU-13, the antibody upon which the humanised
CC antibody of the invention is based
XX
SQ Sequence 360 BP; 90 A; 92 C; 99 G; 79 T; 0 U; 0 Other;
Query Match 81.2%; Score 41.4; DB 6; Length 360;
Best Local Similarity 88.2%; Pred. No. 6.7e-06;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 AAGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 148 AGGATTGATCTCGAATGGTAACTAAATGTAATGACCCGAAGTTCAGGCC 198

RESULT 25
ADF70828
ID ADF70828 standard; cDNA; 360 BP.
XX
AC ADF70828;
XX
DT 12-FEB-2004 (first entry)
XX
DE NR-LU-13 antibody light chain Fv region cDNA.
XX
KW NR-LU-13; light chain variable region; Fv region; gene; ss;
KW humanised antibody; IGG; N-link glycosylation; cancer; cytostatic.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
XX CDS 1..360
XX FT /*tag= a
XX FT /product= "NR-LU-13 light chain Fv region"
XX FT /partial
XX FT /note= "No start or stop codon shown"
XX
XX US2003119078-A1.
XX
XX 26-JUN-2003.
XX
XX 24-JAN-2002; 2002US-00056794.
XX
XX 07-JUN-1996; 96US-00660362.
XX
XX 09-JUN-1997; 97US-00871488.
XX
XX (NEOR-) NEORX CORP.
XX
XX Graves SS, Reno JM, Mallett RW, Hylarides MD, Searle SMJ;
XX PI Henry AH, Pedersen JT, Rees AR;
XX
XX WPI; 2004-009082/01.
XX
XX P-PSDB; ADF70810.

```



```
RESULT 28
AAT38600
ID AAT38600 standard; cDNA; 409 BP.
XX
XX AC AAT38600;
XX
XX DT 04-DEC-1996 (first entry)
XX
XX DE Chimaeric human/murine MAB ONS-M21 variable heavy region cDNA.
XX
XX KW Heavy; variable region; murine; human; myeloblastoma; chimaera;
XX KW monoclonal antibody; chimera; single stranded Fv region;
XX KW low human antigenicity; diagnosis; treatment; cerebral tumour; reshaped;
XX KW ds.
XX
XX OS Synthetic.
XX
XX PH Key Location/Qualifiers
XX FT sig_peptide 1..57
XX FT mat_peptide /*tag= a
XX FT /*tag= b
XX
XX PN JP08169900-A.
XX
XX PD 02-JUL-1996.
XX
XX PF 18-NOV-1994; 94JP-00285057.
XX
XX PR 19-NOV-1993; 93JP-00291078.
XX PR 18-OCT-1994; 94JP-00252166.
XX
XX PA (CHUS ) CHUGAI PHARM CO LTD.
XX
XX WPI; 1996-358509/36.
XX P-PSDB; AAW04379.
XX
XX Reshaped anti-human myeloblastoma cell human antibody - has low human
XX antigenicity, and is therefore useful for diagnosis and treatment of
XX cerebral tumours, e.g. myeloblastoma.
XX
XX PS Example 2; Page 22; 45pp; Japanese.
XX
XX CC The present sequence encodes the variable heavy region of the chimaeric
XX human/murine monoclonal antibody (MAB) ONS-M21. The MAB was prepd. by
XX combining light and heavy variable region DNA, from a murine anti-human
XX myeloblastoma cell MAB, with human light and heavy constant region
XX sequences, respectively to produce chimaeric human/murine light and heavy
XX chain DNA mols. . A recombinant vector for the expression of the heavy and
XX light chain DNA mols. was prepd., and used to transform a host cell. The
XX host cell was then cultured, and the expression prods. of the heavy and
XX light chain DNA mols. sepd. and connected with a peptide linker to
XX produce a single stranded Fv region. The reshaped Fv region has low human
XX antigenicity, and is therefore expected to be useful as an agent for the
XX diagnosis and treatment of cerebral tumours, e.g. myeloblastoma
XX
XX SQ Sequence 409 BP; 104 A; 105 C; 109 G; 91 T; 0 U; 0 Other;
XX
XX Query Match 81.2%; Score 41.4; DB 2; Length 409;
XX Best Local Similarity 88.2%; Pred. No. 6.8e-06;
XX Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX Qy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 205 AGGATTGATCTCGGATGGTAAATACTAAATATGACCCGAGTTCCAGGCC 255
XX
XX RESULT 29
AAT38653
ID AAT38653 standard; cDNA; 409 BP.
XX
XX AC AAT38653;
XX
XX DT 25-MAR-2003 (revised)
XX DT 24-MAR-1994 (first entry)
XX
XX DE V heavy chain gene of recombinant anti-feline calicivirus antibody.
XX
XX KW Antibody; feline calicivirus; FCV; neutralise; heavy chain; virus;
```

```
XX
XX DT 09-DEC-1996 (first entry)
XX
XX DE Chimaeric human/murine MAB ONS-M21 fragment HEP-RVL-M21-g(gamma)1.
XX
XX KW Murine; human; myeloblastoma; chimaera; monoclonal antibody; chimera;
XX KW single stranded Fv region; low human antigenicity; diagnosis; treatment;
XX KW cerebral tumour; reshaped; ds.
XX
XX OS Synthetic.
XX
XX PH Key Location/Qualifiers
XX FT sig_peptide 1..57
XX FT mat_peptide /*tag= a
XX FT /*tag= b
XX
XX PN JP08169900-A.
XX
XX PD 02-JUL-1996.
XX
XX PF 18-NOV-1994; 94JP-00285057.
XX
XX PR 19-NOV-1993; 93JP-00291078.
XX PR 18-OCT-1994; 94JP-00252166.
XX
XX PA (CHUS ) CHUGAI PHARM CO LTD.
XX
XX WPI; 1996-358509/36.
XX P-PSDB; AAW04396.
XX
XX Reshaped anti-human myeloblastoma cell human antibody - has low human
XX antigenicity, and is therefore useful for diagnosis and treatment of
XX cerebral tumours, e.g. myeloblastoma.
XX
XX PS Example 5; Page 38-39; 45pp; Japanese.
XX
XX CC The present sequence is a fragment of the chimaeric human/murine
XX monoclonal antibody (MAB) ONS-M21 cDNA. The MAB was prepd. by combining
XX light and heavy variable region DNA, from a murine anti-human
XX myeloblastoma cell MAB, with human light and heavy constant region
XX sequences, respectively to produce chimaeric human/murine light and heavy
XX chain DNA mols. . A recombinant vector for the expression of the heavy and
XX light chain DNA mols. was prepd., and used to transform a host cell. The
XX host cell was then cultured, and the expression prods. of the heavy and
XX light chain DNA mols. sepd. and connected with a peptide linker to
XX produce a single stranded Fv region. The reshaped Fv region has low human
XX antigenicity, and is therefore expected to be useful as an agent for the
XX diagnosis and treatment of cerebral tumours, e.g. myeloblastoma
XX
XX SQ Sequence 409 BP; 96 A; 106 C; 116 G; 91 T; 0 U; 0 Other;
XX
XX Query Match 81.2%; Score 41.4; DB 2; Length 409;
XX Best Local Similarity 88.2%; Pred. No. 6.8e-06;
XX Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX Qy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 205 AGGATTGATCTCGGATGGTAAATACTAAATATGACCCGAGTTCCAGGCC 255
XX
XX RESULT 30
AAQ48001
ID AAQ48001 standard; DNA; 436 BP.
XX
XX AC AAQ48001;
XX
XX DT 25-MAR-2003 (revised)
XX DT 24-MAR-1994 (first entry)
XX
XX DE V heavy chain gene of recombinant anti-feline calicivirus antibody.
XX
XX KW Antibody; feline calicivirus; FCV; neutralise; heavy chain; virus;
```

```
KW infection; diagnosis; treatment; prophylaxis; ss.
XX Mus musculus.
XX
XX OS
XX
XX Key Location/Qualifiers
XX   16..423
XX CDS /*tag= a
XX   /product= "V heavy chain antibody fragment."
XX
XX
XX EP561194-A2.
XX
XX
XX 22-SEP-1993.
XX
XX 26-FEB-1993; 93EP-00103066.
XX
XX 28-FEB-1992; 92JP-00079189.
XX
XX (KAGA ) CHEMO SERO THERAPEUTIC RES INST.
XX
XX Kimachi K, Maeda H, Nishiyama K, Tokiyoshi S, Tohya Y, Mikami T;
XX WPI; 1993-296521/38.
XX P-PSDB; AAR41679.
XX
XX Anti-feline calicivirus recombinant antibody - used to treat, prevent and
XX diagnosis infection and is safe but effective in cats.
XX
XX Claim 6; Fig 3; 36pp; English.
XX
XX The gene fragment encoding the V heavy chain region is used to produce a
XX genetically engineered antibody capable of reacting specifically with
XX feline calicivirus (FCV). The antibody can be used to treat, prevent and
XX diagnose FCV infection. It has a broad virus- neutralising spectrum and
XX can be used to neutralise viruses which have acquired immunity to other
XX neutralising antibodies. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 436 BP; 107 A; 106 C; 122 G; 101 T; 0 U; 0 Other;
XX
XX Query Match 81.2%; Score 41.4; DB 2; Length 436;
XX Best Local Similarity 88.2%; Pred. No. 6.9e-06;
XX Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCGAGGCC 51
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 220 AGGATTGATCCTCGAATGTTAACTAAATATGACCCGAGTTCCAGGCC 270
XX
XX RESULT 31
XX AAQ99894
XX ID AAQ99894 standard; cDNA; 454 BP.
XX
XX AC AAQ99894;
XX
XX 23-MAR-1996 (first entry)
XX
XX Human VLA-4 reshaped antibody 21.6 heavy chain variable region.
XX
XX Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;
XX antibody engineering; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_signal 7..18
XX /*tag= j
XX /note= "Kozak sequence"
XX CDS 16..441
XX /*tag= a
XX sig_peptide 16..72
XX /*tag= b
XX /*note= "leader"
XX misc_feature 73..162
XX /*tag= c
```

```
FT misc_feature /note= "framework region 1"
FT 163..177
FT /*tag= d
FT /note= "complementarity determining region 1"
FT 178..219
FT /*tag= e
FT /note= "framework region 2"
FT 220..270
FT /*tag= f
FT /note= "complementarity determining region 2"
FT 271..366
FT /*tag= g
FT /note= "framework region 3"
FT 367..408
FT /*tag= h
FT /note= "complementarity determining region 3"
FT 409..441
FT /*tag= i
FT /note= "framework region 4"
FT
XX WO9519790-A1.
XX
XX 27-JUL-1995.
XX
XX 25-JAN-1995; 95WO-US001219.
XX PF
XX 25-JAN-1994; 94US-00186269.
XX PR
XX
XX (ATHE-) ATHENA NEUROSCIENCES INC.
XX
XX Bendig MM, Leger OJ, Saldanha J, Jones TS;
XX WPI; 1995-269276/35.
XX DR P-PSDB; AAR81333.
XX
XX New humanised antibodies against VLA-4 - used for inhibiting leukocyte
XX adhesion to endothelial cells, partic. for treating inflammatory disease.
XX
XX Disclosure; Fig 11; 105pp; English.
XX
XX The sequence encodes the human reshaped antibody 21.6 heavy chain
XX variable region directed against leukocyte adhesion molecule VLA-4.
XX Cloned cDNA sequences of mouse 21.6 VH (AAQ99892) and VL (AAQ99889)
XX regions are linked to human constant regions in the construction of a
XX humanized antibody against VLA-4. The 5' and 3' ends of the mouse cDNAs
XX are modified using PCR primers (See AAQ99895-98) and then subcloned into
XX mammalian cell expression vectors containing human kappa or gamma-1
XX constant regions. In the humanized heavy chain, amino acids H27, H28,
XX H29, H30, H44 and H71 in the human HC VR framework are replaced by the
XX amino acid present in the equivalent position of the mouse 21.6 Ig H
XX chain. Plasmids encoding the chimeric antibodies are transfected into COS
XX cells. The humanized antibodies can be used to inhibit adhesion of a
XX leukocyte to an endothelial cell and for treating inflammatory diseases
XX such as multiple sclerosis. They can also be used in the treatment of
XX stroke, cerebral traumas, meningitis or encephalitis. The antibodies can
XX also be used for detecting VLA-4, for affinity purification or for
XX generating anti-idiotypic antibodies
XX
XX Sequence 454 BP; 104 A; 127 C; 129 G; 94 T; 0 U; 0 Other;
XX
```

```
Query Match 81.2%; Score 41.4; DB 2; Length 454;
Best Local Similarity 88.2%; Pred. No. 6.9e-06;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 220 AGGATTGATCCTCGAATGTTAACTAAATATGACCCGAGTTCCAGGCC 270
```

```
RESULT 32
AAT74789
ID AAT74789 standard; cDNA; 454 BP.
XX
```

AC AAT74789;  
 XX  
 DT 09-DEC-1997 (first entry)  
 XX  
 DE Alpha-4 integrin humanised antibody 21.6 VH version 'a' cDNA.  
 XX  
 KW Alpha-4 integrin; humanised antibody; monoclonal antibody 21.6; asthma;  
 KW atherosclerosis; AIDS; dementia; diabetes; tumour; metastasis;  
 KW inflammatory bowel disease; rheumatoid arthritis; transplant rejection;  
 KW graft versus host disease; nephritis; atopic dermatitis; psoriasis;  
 KW myocardial ischaemia; acute leukocyte mediated lung injury; therapy; ds.  
 XX  
 OS Mus; musculus.  
 OS Homo; sapiens.  
 OS Synthetic.  
 OS Chimeric.  
 XX  
 XX Key Location/Qualifiers  
 FT sig\_peptide 16..72 /\*tag= a  
 FT mat\_peptide 73..441 /\*tag= b  
 FT  
 XX W09718838-A1.  
 XX  
 XX 29-MAY-1997.  
 XX  
 XX 21-NOV-1996; 96WO-US018807.  
 XX  
 XX 21-NOV-1995; 95US-00561521.  
 XX (ATHE-) ATHENA NEUROSCIENCES INC.  
 XX  
 XX Bendig MM, Leger CJ, Saldanha J, Jones ST, Yednock TA;  
 XX WPI: 1997-297879/27.  
 XX P-PSDB; AAW22428.  
 XX  
 XX Uses of humanised alpha-4 integrin antibody - for treatment of asthma,  
 XX atherosclerosis, AIDS, dementia, etc.  
 XX  
 XX Example 6; Fig 11; 107pp; English.  
 XX  
 XX This cDNA clone, designated pUC-res21.6VHa, codes for version 'a'  
 XX (AAW22428) of a reshaped human alpha-4 integrin antibody heavy chain  
 XX variable region (VH) in which complementarity determining regions of the  
 XX alpha-4 integrin mouse monoclonal antibody 21.6 VH region (see AAW22410)  
 XX are incorporated into a human modified 21/28'CL framework (see AAW22413).  
 XX It was constructed from overlapping PCR fragments (see AAT74774-78) and  
 XX can be incorporated into a vector to allow production of reshaped human  
 XX 21.6 VH in transfected mammalian host cells. Claimed humanised 21.6  
 XX antibodies are useful in the manufacture of medicaments for treating  
 XX atherosclerosis, AIDS, dementia, diabetes, inflammatory bowel disease,  
 XX rheumatoid arthritis, transplant rejection, graft versus host disease,  
 XX tumour metastasis, nephritis, atopic dermatitis, psoriasis, myocardial  
 XX ischaemia, asthma and acute leukocyte mediated lung injury  
 XX  
 SQ Sequence 454 BP; 104 A; 127 C; 129 G; 94 T; 0 U; 0 Other;  
 Query Match 81.2%; Score 41.4; DB 2; Length 454;  
 Best Local Similarity 88.2%; Pred. No. 6.9e-06;  
 Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1 AAGATTGATCTCGGAATGGTAAATATATACCCGATATTCAGGCC 51  
 |||||  
 Db 220 AGGATTGATCTCGGAATGGTAAATATATACCCGAGTTCCAGGCC 270  
 |||||  
 RESULT 33.  
 ADR40431  
 ID ADR40431 standard; cDNA; 454 BP.  
 XX  
 AC ADR40431;

XX  
 DT 04-NOV-2004 (first entry)  
 XX  
 DE Reshaped human (RH) 21.6 antibody VH region version a cDNA sequence.  
 XX  
 KW Sulphonylamine; remyelination; demyelinating disease;  
 KW lymphocyte infiltration; paralysis; neuroprotective; muscular;  
 KW cystostatic; antiinflammatory; vulnery; SAMI; antibody; 21.6; human; gene;  
 KW selective adhesion molecule inhibitor;  
 XX ss.  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 16..441  
 FT /\*tag= b  
 FT /product= "Vh region"  
 FT /partial  
 FT /note= "the stop codon is not indicated"  
 FT sig\_peptide 16..72  
 FT /\*tag= a  
 FT /note= "leader sequence"  
 FT  
 XX W02004066931-A2.  
 XX  
 XX 12-AUG-2004.  
 XX  
 XX 26-JAN-2004; 2004WO-US002028.  
 XX  
 XX 24-JAN-2003; 2003US-0442171P.  
 XX 05-SEP-2003; 2003US-0500316P.  
 XX (ELAN-) ELAN PHARM INC.  
 XX  
 XX Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;  
 XX Messersmith E, Freedman S, Yednock T;  
 XX WPI: 2004-593990/57.  
 XX P-PSDB; ADR40432.  
 XX  
 XX Use of sulfonylamine compounds as selective adhesion molecule inhibitors  
 XX to treat demyelinating disease e.g. multiple sclerosis, by inhibiting  
 XX lymphocyte infiltration and promoting remyelination to treat paralysis.  
 XX  
 XX Example 3; Fig 16A-B; 573pp; English.  
 XX  
 XX The invention relates to sulphonylamine compounds (A) of specified  
 XX formulae. The compounds are used to prepare a medicament to promote  
 XX remyelination of nerve cells in a mammal for treatment of a demyelinating  
 XX disease and to inhibit lymphocyte infiltration of immune cells in the  
 XX spinal cord to promote remyelination of nerve cells in the spinal cord,  
 XX thus treating paralysis. The method of inhibiting lymphocyte infiltration  
 XX of immune cells in the spinal cord to promote remyelination of nerve  
 XX cells in the spinal cord and thus treating paralysis further comprises co  
 XX -administering an immunosuppressant such as adrenocorticotrophic hormone,  
 XX a corticosteroid (e.g. prednisone, methylprednisolone, dexamethasone  
 XX cortisol, cortisone, fludrocortisone or betamethasone) or an interferon  
 XX (e.g. interferon beta-1b or interferon beta-1a). (A) is used to prepare a  
 XX medicament for treating paralysis, a congenital metabolic disorder, a  
 XX neuropathy with abnormal myelination, drug induced demyelination,  
 XX radiation induced demyelination, hereditary demyelinating condition,  
 XX prior induced demyelinating condition, encephalitis induced demyelination  
 XX or spinal cord injury. The present sequence represents a cDNA encoding  
 XX the first version (a) of a reshaped human 21.6 antibody heavy chain  
 XX variable region.  
 XX  
 SQ Sequence 454 BP; 104 A; 127 C; 129 G; 94 T; 0 U; 0 Other;  
 Query Match 81.2%; Score 41.4; DB 13; Length 454;  
 Best Local Similarity 88.2%; Pred. No. 6.9e-06;  
 Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;





XX Disclosure; Fig 2; 105pp; English.

XX The sequence encodes the mouse antibody 21.6 heavy chain variable region

CC directed against leukocyte adhesion molecule VLA-4. Cloned cDNA sequences

CC of mouse 21.6 VH and VL (see AAQ99889) regions are linked to human

CC constant regions in the construction of a humanized antibody against VLA-

CC 4. The 5' and 3' ends of the mouse cDNAs are modified using PCR primers

CC (See AAQ99893-98) and then subcloned into mammalian cell expression

CC vectors containing human kappa or gamma-1 constant regions. In the

CC humanized heavy chain, amino acids H27, H28, H29, H30, H44 and H71 in the

CC human HC VR framework are replaced by the amino acid present in the

CC equivalent position of the mouse 21.6 Ig H chain. Plasmids encoding the

CC chimeric antibodies are transfected into COS cells. The humanized

CC antibodies can be used to inhibit adhesion of a leukocyte to an

CC endothelial cell and for treating inflammatory diseases such as multiple

CC sclerosis. They can also be used in the treatment of stroke, cerebral

CC trauma, meningitis or encephalitis. The antibodies can also be used for

CC detecting VLA-4, for affinity purification or for generating anti-

CC idotype antibodies

XX SQ Sequence 470 BP; 118 A; 121 C; 123 G; 108 T; 0 U; 0 Other;

Query Match 81.2%; Score 41.4; DB 2; Length 470;

Best Local Similarity 88.2%; Pred. No. 6.9e-06;

Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51

Db 205 AGGATTGATCTCGGAATGGTTATCTACTAAATATGACCCGAGTTCCAGGCC 255

RESULT 36

AAAT4760

ID AAT4760 standard; DNA; 470 BP.

XX AC AAT4760;

XX DT 08-DEC-1997 (first entry)

XX DE Alpha-4 integrin mouse MAb 21.6 VH region DNA.

KW Alpha-4 integrin; humanised antibody; monoclonal antibody 21.6; asthma;

KW atherosclerosis; AIDS; dementia; diabetes; tumour; metastasis;

KW inflammatory bowel disease; rheumatoid arthritis; transplant rejection;

KW graft versus host disease; nephritis; atopic dermatitis; psoriasis;

KW myocardial ischaemia; acute leukocyte mediated lung injury; therapy; ds.

OS Mus musculus.

XX FH Key Location/Qualifiers

FT sig\_peptide 1..57

FT mat\_peptide 58..420

FT /\*tag= a

FT /\*tag= b

XX W09718838-A1.

XX DT 29-MAY-1997.

XX DT 21-NOV-1996; 96WO-US018807.

XX DT 21-NOV-1995; 95US-00561521.

XX (ATHE-) ATHENA NEUROSCIENCES INC.

XX Bendig MM, Leger OJ, Saldanha J, Jones ST, Yednock TA;

XX WPI; 1997-297879/27.

DR P-PSDB; AAW22410.

XX Uses of humanised alpha-4 integrin antibody - for treatment of asthma,

PT atherosclerosis, AIDS, dementia, etc.

XX Example 1; Fig 2; 107pp; English.

XX This DNA sequence codes for the heavy chain variable region (VH)

CC (AAW22410) of mouse anti-alpha-4 integrin monoclonal antibody 21.6. It

CC was produced from 21.6 hybridoma cDNA by PCR amplification. The DNA can

CC be used in the preparation of a claimed humanised 21.6 antibody VH region

CC (see AAW22413) in which complementarity determining regions of the mouse

CC 21.6 VL are incorporated into a human 21/28'CL framework. Claimed

CC humanised 21.6 antibodies are useful in the manufacture of medicaments

CC for treating asthma, atherosclerosis, AIDS, dementia, diabetes,

CC inflammatory bowel disease, rheumatoid arthritis, transplant rejection,

CC graft versus host disease, tumour metastasis, nephritis, atopic

CC dermatitis, psoriasis, myocardial ischaemia and acute leukocyte mediated

CC lung injury. The humanised antibodies have a half-life in the human

CC circulation essentially equivalent to that of natural human antibodies

XX SQ Sequence 470 BP; 118 A; 121 C; 123 G; 108 T; 0 U; 0 Other;

Query Match 81.2%; Score 41.4; DB 2; Length 470;

Best Local Similarity 88.2%; Pred. No. 6.9e-06;

Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51

Db 205 AGGATTGATCTCGGAATGGTTATCTACTAAATATGACCCGAGTTCCAGGCC 255

RESULT 37

ADR40418

ID ADR40418 standard; DNA; 470 BP.

XX AC ADR40418;

XX DT 04-NOV-2004 (first entry)

XX DE Mouse 21.6 antibody Vh region DNA sequence.

KW Sulphonylamine; remyelination; demyelinating disease;

KW lymphocyte infiltration; paralysis; neuroprotective; muscular;

KW cytostatic; antiinflammatory; vulnery;

KW selective adhesion molecule inhibitor; SAMI; antibody; 21.6; mouse; gene;

ds.

XX OS Mus sp.

XX FH Key Location/Qualifiers

FT CDS 1..432

FT /\*tag= b

FT /product= "Vh region"

FT /partial

FT /note= "the stop codon is not indicated"

FT /transl\_except= pos: 427..429, aa: Ser)

FT /transl\_except= pos: 430..432, aa: Ser)

FT sig\_peptide 1..57

FT /\*tag= a

FT /note= "leader sequence"

XX W02004066931-A2.

XX DT 12-AUG-2004.

XX DT 26-JAN-2004; 2004WO-US002028.

XX DT 24-JAN-2003; 2003US-0442171P.

XX DT 05-SEP-2003; 2003US-0500316P.

XX (ELAN-) ELAN PHARM INC.

XX Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;

PI Messersmith E, Freedman S, Yednock T;

XX WPI; 2004-593990/57.

```

DR P-PSDB; ADR40419.
XX
XX Use of sulfonamide compounds as selective adhesion molecule inhibitors
XX to treat demyelinating disease e.g. multiple sclerosis, by inhibiting
XX lymphocyte infiltration and promoting remyelination to treat paralysis.
XX
XX Example 3; Fig 12A-B; 573pp; English.
XX
XX The invention relates to sulphonylamine compounds (A) of specified
XX formulae. The compounds are used to prepare a medicament to promote
XX remyelination of nerve cells in a mammal for treatment of a demyelinating
XX disease and to inhibit lymphocyte infiltration of immune cells in the
XX spinal cord to promote remyelination of nerve cells in the spinal cord,
XX thus treating paralysis. The method of inhibiting lymphocyte infiltration
XX of immune cells in the spinal cord to promote remyelination of nerve
XX cells in the spinal cord and thus treating paralysis further comprises co
XX administering an immunosuppressant such as adrenocorticotrophic hormone,
XX a corticosteroid (e.g. prednisone, methylprednisolone, dexamethasone
XX cortisone, fludrocortisone, prednisolone, alpha-
XX methylprednisolone, triamcinolone or betamethasone) or an interferon
XX (e.g. interferon beta-1b or interferon beta-1a). (A) is used to prepare a
XX medicament for treating paralysis, a congenital metabolic disorder, a
XX neuropathy with abnormal myelination, drug induced demyelination,
XX radiation induced demyelination, hereditary demyelinating condition,
XX prion induced demyelinating condition, encephalitis induced demyelination
XX or spinal cord injury. The present sequence represents a mouse 21.6
XX antibody heavy chain variable region (Vh) DNA sequence.
XX
XX Sequence 470 BP; 118 A; 121 C; 123 G; 108 T; 0 U; 0 Other;
XX
XX Query Match 81.2%; Score 41.4; DB 13; Length 470;
XX Best Local Similarity 88.2%; Pred. No. 6.9e-06;
XX Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX 1 AAGATTGATCTCGAATGTTAACTAAATATGACCCGATATTCAGGCC 51
XX 205 AGGATTGATCTCGAATGTTAACTAAATATGACCCGATATTCAGGCC 255
XX
XX RESULT 39
XX ADR40336
XX ID ADR40336 standard; DNA; 470 BP.
XX
XX AC ADR40336;
XX
XX 04-NOV-2004 (first entry)
XX
XX Mouse 21.6 antibody Vh region DNA sequence.
XX
XX Remyelination; demyelinating disease; lymphocyte infiltration; paralysis;
XX neuroprotective; muscular; cytostatic; antiinflammatory; vulnary;
XX selective adhesion molecule inhibitor; SAMI; antibody; 21.6; mouse; gene;
XX ds.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX CDS 1..432
XX /*tag= b
XX /product= "Vh region"
XX /partial
XX /note= "the stop codon is not indicated"
XX /transl_except= pos: 427..429, aa: Ser)
XX /transl_except= pos: 430..432, aa: Ser)
XX
XX sig_peptide 1..57
XX /*tag= a
XX /note= "leader sequence"
XX
XX WO2004066932-A2.
XX
XX 12-AUG-2004.
XX
XX 26-JAN-2004; 2004WO-US002039.
XX
XX 24-JAN-2003; 2003US-0442171P.
XX 05-SEP-2003; 2003US-0500316P.
XX
XX (ELAN-) ELAN PHARM INC.
XX
XX Karlik SJ, Pleiss MA, Konradi AW, Grant FS, Semko CM, Dressen D;
XX Messersmith E, Freedman S, Yednock T;
XX
XX WPI; 2004-593991/57.
XX P-PSDB; ADR40337.
XX
XX Use of remyelinating agent such as natalizumab for preparation of
XX medicament for treating demyelinating diseases and paralysis by promoting
XX remyelination of nerve cells in mammal.
XX
XX Example 3; Fig 12A-B; 495pp; English.
XX
XX The invention relates to remyelinating agents (A) that are used to
XX prepare a medicament to promote remyelination of nerve cells in a mammal
XX for treatment of a demyelinating disease and to inhibit lymphocyte
XX infiltration of immune cells in the spinal cord to promote remyelination
XX of nerve cells in the spinal cord, thus treating paralysis. The method of
XX inhibiting lymphocyte infiltration of immune cells in the spinal cord to
XX promote remyelination of nerve cells in the spinal cord and thus treating
XX paralysis further comprises co-administering an immunosuppressant such as
XX adrenocorticotrophic hormone, a corticosteroid (e.g. prednisone,
XX methylprednisolone, dexamethasone cortisone, cortisone, fludrocortisone,
XX prednisolone, alpha-methylprednisolone, triamcinolone or betamethasone) or
XX an interferon (e.g. interferon beta-1b or interferon beta-1a). (A) is
XX used to prepare a medicament for treating paralysis, a congenital
XX metabolic disorder, a neuropathy with abnormal myelination, drug induced
XX demyelination, radiation induced demyelination, hereditary demyelinating
XX condition, prion induced demyelinating condition, encephalitis induced
XX demyelination or spinal cord injury. The present sequence represents a
XX mouse 21.6 antibody heavy chain variable region (Vh) DNA sequence.
XX
XX Sequence 470 BP; 118 A; 121 C; 123 G; 108 T; 0 U; 0 Other;
XX
XX Query Match 81.2%; Score 41.4; DB 13; Length 470;
XX Best Local Similarity 88.2%; Pred. No. 6.9e-06;
XX Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX 1 AAGATTGATCTCGAATGTTAACTAAATATGACCCGATATTCAGGCC 51
XX 205 AGGATTGATCTCGAATGTTAACTAAATATGACCCGATATTCAGGCC 255
XX
XX RESULT 39
XX AAQ94548
XX ID AAQ94548 standard; cDNA; 822 BP.
XX
XX AC AAQ94548;
XX
XX 18-JAN-1996 (first entry)
XX
XX Plasmid pSCFVT7-hm21.
XX
XX Plasmid pSCFVT7-hm21; human; ONS-M21 antibody; chimeric protein;
XX medulloblastoma; brain tumour; treatment; diagnosis; Fv fragment; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..810
XX /*tag= a
XX
XX sig_peptide 1..66
XX /*tag= b
XX
XX mat_peptide 67..807
XX /*tag= c
XX
XX WO9514041-A1.
XX

```

PD 26-MAY-1995.  
 XX  
 PF 19-OCT-1994; 94WO-JP001763.  
 XX  
 PR 19-NOV-1993; 93JP-00291078.  
 XX  
 PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX  
 PI Ohtomo T, Sato K, Tsuchiya M;  
 XX WPI; 1995-200347/26.  
 DR P-PSDB; AAR76682.  
 XX  
 PT Reconstituted antibody against human medullo:blastoma cells - contains  
 PT high proportion of human antibody origin and has low antigenicity.  
 XX  
 PS Claim 35; Page 98-99; 120pp; Japanese.  
 XX  
 CC AAQ94548 is the plasmid pSCFV7-bm21, which encodes AAR76682 the human  
 CC antibody ONS-M21 Fv fragment. The plasmid was used in the construction of  
 CC an expression vector, con9g. cDNA encoding a human/murine chimeric  
 CC antibody, reactive with human medullo- blastoma (a brain tumour) cells.  
 CC The chimeric antibody can be used in the diagnosis and treatment of this  
 CC disease  
 XX  
 SQ Sequence 822 BP; 210 A; 212 C; 232 G; 168 T; 0 U; 0 Other;  
 Query Match 81.2%; Score 41.4; DB 2; Length 822;  
 Best Local Similarity 88.2%; Pred. No. 7.6e-06;  
 Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
 DB 214 AGGATTGATCCTCGGATGTTAACTAAATATGACCCGATATTCAGGCC 264  
 Search completed: June 4, 2005, 13:53:54  
 Job time : 314.242 secs

---

PT antigenicity, and is therefore useful for diagnosis and treatment of  
 PT cerebral tumours, e.g. myeloblastoma.  
 XX  
 PS Example 6; Page 40-41; 45pp; Japanese.  
 XX  
 CC The present sequence is a scFv fragment from the chimaeric human/murine  
 CC monoclonal antibody (MAB) ONS-M21 cDNA. The MAB was prepd. by combining  
 CC light and heavy variable region DNA, from a murine anti-human  
 CC myeloblastoma cell MAB, with human light and heavy constant region  
 CC sequences, respectively to produce chimaeric human/murine light and heavy  
 CC chain DNA mols. . A recombinant vector for the expression of the heavy and  
 CC light chain DNA mols. was prepd., and used to transform a host cell. The  
 CC host cell was then cultured, and the expression prods. of the heavy and  
 CC light chain DNA mols. sepd. and connected with a peptide linker to  
 CC produce a single stranded Fv region. The reshaped Fv region has low human  
 CC antigenicity, and is therefore expected to be useful as an agent for the  
 CC diagnosis and treatment of cerebral tumours, e.g. myeloblastoma  
 XX  
 SQ Sequence 822 BP; 210 A; 213 C; 232 G; 167 T; 0 U; 0 Other;  
 Query Match 81.2%; Score 41.4; DB 2; Length 822;  
 Best Local Similarity 88.2%; Pred. No. 7.6e-06;  
 Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
 DB 214 AGGATTGATCCTCGGATGTTAACTAAATATGACCCGATATTCAGGCC 264  
 Search completed: June 4, 2005, 13:53:54  
 Job time : 314.242 secs

---

PD 26-MAY-1995.  
 XX  
 PF 19-OCT-1994; 94WO-JP001763.  
 XX  
 PR 19-NOV-1993; 93JP-00291078.  
 XX  
 PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX  
 PI Ohtomo T, Sato K, Tsuchiya M;  
 XX WPI; 1995-200347/26.  
 DR P-PSDB; AAR76682.  
 XX  
 PT Reconstituted antibody against human medullo:blastoma cells - contains  
 PT high proportion of human antibody origin and has low antigenicity.  
 XX  
 PS Claim 35; Page 98-99; 120pp; Japanese.  
 XX  
 CC AAQ94548 is the plasmid pSCFV7-bm21, which encodes AAR76682 the human  
 CC antibody ONS-M21 Fv fragment. The plasmid was used in the construction of  
 CC an expression vector, con9g. cDNA encoding a human/murine chimeric  
 CC antibody, reactive with human medullo- blastoma (a brain tumour) cells.  
 CC The chimeric antibody can be used in the diagnosis and treatment of this  
 CC disease  
 XX  
 SQ Sequence 822 BP; 210 A; 212 C; 232 G; 168 T; 0 U; 0 Other;  
 Query Match 81.2%; Score 41.4; DB 2; Length 822;  
 Best Local Similarity 88.2%; Pred. No. 7.6e-06;  
 Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
 DB 214 AGGATTGATCCTCGGATGTTAACTAAATATGACCCGATATTCAGGCC 264  
 Search completed: June 4, 2005, 13:53:54  
 Job time : 314.242 secs

---

RESULT 40  
 AAT38662  
 ID AAT38662 standard; cDNA; 822 BP.  
 XX  
 AC AAT38662;  
 XX  
 DT 09-DEC-1996 (first entry)  
 XX  
 DE Chimaeric human/murine MAB ONS-M21 scFv fragment cDNA.  
 XX  
 KW Murine; human; myeloblastoma; chimera; monoclonal antibody; chimera;  
 KW single stranded Fv region; low human antigenicity; diagnosis; treatment;  
 KW cerebral tumour; reshaped; ds.  
 XX  
 OS Synthetic.  
 XX  
 PH Key Location/Qualifiers  
 PH CDS 1..810  
 FT /\*tag= a  
 FT sig\_peptide 1..66  
 FT /\*tag= a  
 XX  
 PN JP08169900-A.  
 XX  
 PD 02-JUL-1996.  
 XX  
 PF 18-NOV-1994; 94JP-00285057.  
 XX  
 PR 19-NOV-1993; 93JP-00291078.  
 PR 18-OCT-1994; 94JP-00252166.  
 XX  
 PA (CHUS ) CHUGAI PHARM CO LTD.  
 XX  
 WPI; 1996-358509/36.  
 DR P-PSDB; AAW04397.  
 XX  
 PT Reshaped anti-human myeloblastoma cell human antibody - has low human

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 13:25:28 ; Search time 100.455 Seconds  
(without alignment)  
830.725 Million cell updates/sec

Title: US-10-089-452-25

Perfect score: 51

Sequence: 1 aagattgactctgcgaatgg.....atgaccgcgatattccaggcc 51

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	84.3	412	3	US-08-836-561-30
2	43	84.3	412	4	Sequence 30, Appl
3	41.6	81.6	339	1	US-09-434-122-30
C 4	41.4	81.2	132	3	Sequence 1, Appl
5	41.4	81.2	329	2	US-08-207-169A-1
6	41.4	81.2	329	2	US-08-646-285A-95
7	41.4	81.2	360	3	Sequence 95, Appl
8	41.4	81.2	409	1	Sequence 3, Appl
9	41.4	81.2	409	3	Sequence 14, Appl
10	41.4	81.2	409	3	Sequence 1, Appl
11	41.4	81.2	454	2	Sequence 28, Appl
12	41.4	81.2	454	5	Sequence 98, Appl
13	41.4	81.2	470	2	Sequence 16, Appl
14	41.4	81.2	470	5	PCT-US95-01219-16
15	41.4	81.2	822	3	Sequence 3, Appl
16	39.8	78.0	357	4	Sequence 108, App
17	39.8	78.0	414	4	Sequence 19, Appl
18	39.8	78.0	1632	2	Sequence 21, Appl
19	39.8	78.0	1641	2	Sequence 8, Appl
20	39.8	78.0	1644	2	Sequence 5, Appl
21	39.8	78.0	1672	2	Sequence 11, Appl
22	39.8	78.0	4435	2	Sequence 2, Appl
23	38.4	75.3	354	2	Sequence 1, Appl
24	38.4	75.3	354	2	Sequence 37, Appl
25	38.2	74.9	71	4	Sequence 19, Appl
26	38.2	74.9	360	1	Sequence 1, Appl
27	38.2	74.9	360	1	Sequence 1, Appl

US-08-836-561-30  
; Sequence 30, Application US/08836561  
; Patent No. 6018032  
; GENERAL INFORMATION:  
; APPLICANT: KOIKE, Masamichi  
; APPLICANT: FURUYA, Akiko  
; APPLICANT: NAKAMURA, Kazuyasu  
; APPLICANT: IIDA, Akihiro  
; APPLICANT: ANAZAWA, Hideharu  
; APPLICANT: HANAI, No. 6018032uo  
; APPLICANT: TAKATSU, Kiyoshi  
; TITLE OF INVENTION: Antibody Against Human Interleukin-5  
; TITLE OF INVENTION: Receptor Alpha Chain  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/836,561  
; FILING DATE: 09-MAY-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 232384/95  
; FILING DATE: 11-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lawrence, III, Stanton T  
; REGISTRATION NUMBER: 25,736  
; REFERENCE/DOCKET NUMBER: 7005-115-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 412 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-836-561-30

#### ALIGNMENTS

RESULT 1  
US-08-836-561-30  
; Sequence 30, Application US/08836561  
; Patent No. 6018032  
; GENERAL INFORMATION:  
; APPLICANT: KOIKE, Masamichi  
; APPLICANT: FURUYA, Akiko  
; APPLICANT: NAKAMURA, Kazuyasu  
; APPLICANT: IIDA, Akihiro  
; APPLICANT: ANAZAWA, Hideharu  
; APPLICANT: HANAI, No. 6018032uo  
; APPLICANT: TAKATSU, Kiyoshi  
; TITLE OF INVENTION: Antibody Against Human Interleukin-5  
; TITLE OF INVENTION: Receptor Alpha Chain  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/836,561  
; FILING DATE: 09-MAY-1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 232384/95  
; FILING DATE: 11-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lawrence, III, Stanton T  
; REGISTRATION NUMBER: 25,736  
; REFERENCE/DOCKET NUMBER: 7005-115-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 412 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-836-561-30

28 38.2 74.9 360 2 US-08-436-339A-1 Sequence 1, Appli  
29 38.2 74.9 360 2 US-08-950-660-1 Sequence 1, Appli  
30 38.2 74.9 360 4 US-08-454-899G-5 Sequence 5, Appli  
31 38.2 74.9 360 5 PCT-US93-00030-1 Sequence 1, Appli  
32 38.2 74.9 360 5 PCT-US93-00924-1 Sequence 1, Appli  
33 38.2 74.9 360 5 PCT-US94-10395-1 Sequence 1, Appli  
34 38.2 74.9 363 2 US-08-822-830B-1 Sequence 1, Appli  
35 38.2 74.9 363 2 US-08-822-830B-12 Sequence 12, Appli  
36 38.2 74.9 363 4 US-09-157-452B-1 Sequence 1, Appli  
37 38.2 74.9 372 4 US-08-157-452B-15 Sequence 15, Appli  
38 38.2 74.9 429 1 US-08-454-899G-46 Sequence 46, Appli  
39 38.2 74.9 429 1 US-08-463-128-3 Sequence 3, Appli  
40 38.2 74.9 429 1 US-08-463-128-7 Sequence 7, Appli  
41 38.2 74.9 429 1 US-08-463-298-3 Sequence 3, Appli  
42 38.2 74.9 429 1 US-08-463-298-7 Sequence 7, Appli  
43 38.2 74.9 429 2 US-08-436-339A-3 Sequence 3, Appli  
44 38.2 74.9 429 2 US-08-436-339A-7 Sequence 7, Appli  
45 38.2 74.9 429 4 US-08-454-899G-28 Sequence 28, Appli



```

;
OPERATING SYSTEM: PC-DOS/MS-DOS
;
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

REFERENCE/DOCKET NUMBER: HOAG-1043

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4000  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 329 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-463-087-3

Query Match 81.2%; Score 41.4; DB 2; Length 329;  
Best Local Similarity 88.2%; Pred. No. 6.9e-07;  
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 141 AGGATTGCTCGCAATGTTAACTAAATATGACCCGATATTCAGGCC 191

RESULT 7  
US-08-871-488A-14  
Sequence 14, Application US/08871488A  
Patent No. 6358710  
GENERAL INFORMATION:  
APPLICANT: Graves, Scott S.  
APPLICANT: Reno, John M.  
APPLICANT: Mallett, Robert W.  
APPLICANT: Hylarides, Mark D.  
APPLICANT: Searle, Stephen M.J.  
APPLICANT: Henry, Andrew H.  
APPLICANT: Pedersen, Jan T.  
APPLICANT: Rees, Anthony R.  
TITLE OF INVENTION: HUMANIZED ANTIBODIES THAT BIND TO THE  
TITLE OF INVENTION: ANTIGEN BOUND BY ANTIBODY NR-LU-13 AND THEIR USE IN  
TITLE OF INVENTION: PRETARGETING METHODS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/871,488A  
FILING DATE: 09-JUN-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey Ph.D., Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 690022.527C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 360 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..360  
US-08-871-488A-14

Query Match 81.2%; Score 41.4; DB 3; Length 360;  
Best Local Similarity 88.2%; Pred. No. 7e-07;

Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 148 AGGATTGATCCTCGAATGTTAACTAAATATGACCCGATATTCAGGCC 198

RESULT 8  
US-08-024-253-1  
Sequence 1, Application US/08024253  
Patent No. 5785968  
GENERAL INFORMATION:  
APPLICANT: KIMACHI, Kazuhiko  
APPLICANT: MAEDA, Hiroaki  
APPLICANT: NISHIYAMA, Kiyoto  
APPLICANT: TOKIYOSHI, Sachio  
APPLICANT: TOHYA, Yukinobu  
APPLICANT: MIKAMI, Takeshi  
TITLE OF INVENTION: ANTI-FELINE CALCIVIRUS RECOMBINANT  
TITLE OF INVENTION: ANTIBODY AND GENE FRAGMENT ENCODING THE SAME  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WEGNER, CANTOR, MUELLER & PLAYER  
STREET: 1233 20th Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20036-8218  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/024,253  
FILING DATE: 19930301  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 79189/1992  
FILING DATE: 28-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: CANTOR, Herbert I.  
REGISTRATION NUMBER: 24,392  
REFERENCE/DOCKET NUMBER: P-500-23744  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-0400  
TELEFAX: (202) 835-0605  
TELEX: 440706 WEGBR  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA to mRNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..408  
US-08-024-253-1

Query Match 81.2%; Score 41.4; DB 1; Length 409;  
Best Local Similarity 88.2%; Pred. No. 7.2e-07;  
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 205 AGGATTGATCCTCGAATGTTAACTAAATATGACCCGATATTCAGGCC 255

RESULT 9  
US-08-646-265A-28  
Sequence 28, Application US/08646265A  
Patent No. 6214973



GENERAL INFORMATION:  
APPLICANT: OHTOMO, Toshihiko  
APPLICANT: SATO, Koh  
APPLICANT: TSUCHIYA, Masayuki  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
TITLE OF INVENTION: MEDULLOBLASTOMA CELLS  
NUMBER OF SEQUENCES: 132  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,265A  
FILING DATE: 09-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/JP94/01763  
FILING DATE: 19-OCT-1994  
APPLICATION NUMBER: JP 5-291078  
FILING DATE: 19-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/184  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..408  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1..408  
US-08-646-265A-28

Query Match 81.2%; Score 41.4; DB 3; Length 409;  
Best Local Similarity 88.2%; Pred. No. 7.2e-07;  
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 AGAGTTGATCTCGGATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 205 AGGATTGATCTCGGATGTTAACTAAATATGACCCGATATTCAGGCC 255

RESULT 10  
US-08-646-265A-98  
Sequence 98, Application US/08646265A  
Patent No. 6214973  
GENERAL INFORMATION:  
APPLICANT: OHTOMO, Toshihiko  
APPLICANT: SATO, Koh  
APPLICANT: TSUCHIYA, Masayuki  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
TITLE OF INVENTION: MEDULLOBLASTOMA CELLS  
NUMBER OF SEQUENCES: 132  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,265A  
FILING DATE: 09-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/JP94/01763  
FILING DATE: 19-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 5-291078  
FILING DATE: 19-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/184  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 98:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..408  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 58..408  
US-08-646-265A-98

Query Match 81.2%; Score 41.4; DB 3; Length 409;  
Best Local Similarity 88.2%; Pred. No. 7.2e-07;  
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AGAGTTGATCTCGGATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 205 AGGATTGATCTCGGATGTTAACTAAATATGACCCGATATTCAGGCC 255

RESULT 11  
US-08-561-521-16  
Sequence 16, Application US/08561521  
Patent No. 5840299  
GENERAL INFORMATION:  
APPLICANT: Bendig, Mary M.  
APPLICANT: Leger, Olivier J.  
APPLICANT: Saldanha, Jose  
APPLICANT: Jones, S. Tarran  
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
TITLE OF INVENTION: Adhesion Molecule VLA-4  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Hourie and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/561,521  
FILING DATE: 25-JAN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/186,269A  
FILING DATE: 25-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William L.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 15270-14  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 454 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 16..441  
PCT-US95-01219-16

Query Match 81.2%; Score 41.4; DB 2; Length 454;  
Best Local Similarity 88.2%; Pred. No. 7.4e-07;  
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 AGATTGATCCTCGAATGTTAACTAAATATGACCCGATATTCACGGCC 51  
Db 220 AGGATTGATCCTCGAATGTTAACTAAATATGACCCGATATTCACGGCC 270

RESULT 12  
PCT-US95-01219-16  
Sequence 16, Application PC/TUS9501219  
GENERAL INFORMATION:  
APPLICANT: Bendig, Mary M.  
APPLICANT: Leger, Olivier J.  
APPLICANT: Saldanha, Jose  
APPLICANT: Jones, S. Tarran  
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
TITLE OF INVENTION: Adhesion Molecule VLA-4  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01219  
FILING DATE: 25-JAN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/186,269  
FILING DATE: 25-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William L.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 15270-14  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600

TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 454 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 16..441  
PCT-US95-01219-16

Query Match 81.2%; Score 41.4; DB 5; Length 454;  
Best Local Similarity 88.2%; Pred. No. 7.4e-07;  
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 AGATTGATCCTCGAATGTTAACTAAATATGACCCGATATTCACGGCC 51  
Db 220 AGGATTGATCCTCGAATGTTAACTAAATATGACCCGATATTCACGGCC 270

RESULT 13  
US-08-561-521-3  
Sequence 3, Application US/08561521  
Patent No. 5840299  
GENERAL INFORMATION:  
APPLICANT: Bendig, Mary M.  
APPLICANT: Leger, Olivier J.  
APPLICANT: Saldanha, Jose  
APPLICANT: Jones, S. Tarran  
TITLE OF INVENTION: Humanized Antibodies Against Leukocyte  
TITLE OF INVENTION: Adhesion Molecule VLA-4  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/561,521  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/186,269A  
FILING DATE: 25-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William L.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 15270-14  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-543-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 470 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..420  
US-08-561-521-3

Query Match 81.2%; Score 41.4; DB 2; Length 470;





TELECOMMUNICATION INFORMATION:  
TELEPHONE: 410-671-1158  
TELEFAX: 410-671-2534  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1632 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: not relevant  
MOLECULE TYPE: cdna  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
STRAIN: BALB/c  
IMMEDIATE SOURCE:  
CLONE: phist 20  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 88..798  
OTHER INFORMATION: /product= "antibody fragment, light  
OTHER INFORMATION: Chain"  
OTHER INFORMATION: /label= BotFab 20  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 818..1582  
OTHER INFORMATION: /product= "antibody fragment, Heavy  
OTHER INFORMATION: Chain"  
OTHER INFORMATION: /label= BotFab 20  
US-08-792-824-8

Query Match 78.0%; Score 39.8; DB 2; Length 1632;  
Best Local Similarity 86.3%; Pred. No. 4.2e-06;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGGAATGTAATACTAAATATGACCCGATATTCACAGGCC 51  
DB 1031 AGATTGATCTCGGAATGTAATACTAAATATGACCCGATATTCACAGGCC 1081

RESULT 19  
US-08-792-824-5  
Sequence 5, Application US/08792824  
Patent No. 5932449  
GENERAL INFORMATION:  
APPLICANT: EMANUEL, PETER A.  
APPLICANT: BURANS, JAMES P.  
APPLICANT: VALDES, JAMES J.  
APPLICANT: MOHVEE, ELDEFRAWI E.  
TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: U.S. Army Chemical and Biological Defense  
ADDRESSEE: Command  
STREET: Office of the Chief Counsel, Bldg E4435  
CITY: Aberdeen Proving Ground  
STATE: MD  
COUNTRY: U.S.  
ZIP: 21010-5423  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/792,824  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Biffoni, U. J.  
REGISTRATION NUMBER: 39,908  
REFERENCE/DOCKET NUMBER: DAM 431-96  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 410-671-1158  
TELEFAX: 410-671-2534  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1641 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: not relevant  
MOLECULE TYPE: cdna  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
STRAIN: BALB/c  
IMMEDIATE SOURCE:  
CLONE: phist 1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 87..788  
OTHER INFORMATION: /product= "antibody fragment, light  
OTHER INFORMATION: Chain"  
OTHER INFORMATION: /label= BotFab 1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 815..1579  
OTHER INFORMATION: /product= "antibody fragment, Heavy  
OTHER INFORMATION: Chain"  
OTHER INFORMATION: /label= BotFab 1  
US-08-792-824-5

Query Match 78.0%; Score 39.8; DB 2; Length 1641;  
Best Local Similarity 86.3%; Pred. No. 4.2e-06;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGGAATGTAATACTAAATATGACCCGATATTCACAGGCC 51  
DB 1028 AGATTGATCTCGGAATGTAATACTAAATATGACCCGATATTCACAGGCC 1078

RESULT 20  
US-08-792-824-11  
Sequence 11, Application US/08792824  
Patent No. 5932449  
GENERAL INFORMATION:  
APPLICANT: EMANUEL, PETER A.  
APPLICANT: BURANS, JAMES P.  
APPLICANT: VALDES, JAMES J.  
APPLICANT: MOHVEE, ELDEFRAWI E.  
TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: U.S. Army Chemical and Biological Defense  
ADDRESSEE: Command  
STREET: Office of the Chief Counsel, Bldg E4435  
CITY: Aberdeen Proving Ground  
STATE: MD  
COUNTRY: U.S.  
ZIP: 21010-5423  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/792,824  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Biffoni, U. J.  
REGISTRATION NUMBER: 39,908  
REFERENCE/DOCKET NUMBER: DAM 431-96  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 410-671-1158



```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 4435 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; STRAIN: BALB/c
; IMMEDIATE SOURCE:
; CLONE: Clone pHist 5
US-08-792-824-1

Query Match 78.0%; Score 39.8; DB 2; Length 4435;
Best Local Similarity 86.3%; Pred. No. 5.3e-06;
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCACGGCC 51
Db 3194 AGGATTGATCTCGGAATGGGATGACTGAATATGACCCGAGTTCCAGGGC 3244

RESULT 23
US-08-232-081B-34
; Sequence 34, Application US/08232081B
; Patent No. 5886152
; GENERAL INFORMATION:
; APPLICANT: NAKATANI, TOMOYUKI
; APPLICANT: GOMI, HIDEYUKI
; APPLICANT: WIJENES, JOHN
; APPLICANT: NOGUCHI, HIROSHI
; TITLE OF INVENTION: HUMANIZED B-B10
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,081B
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 20-3484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 354 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..354
US-08-232-081B-37

Query Match 75.3%; Score 38.4; DB 2; Length 354;
Best Local Similarity 87.5%; Pred. No. 1.1e-05;
Matches 42; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 ATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCACGGCC 51
Db 151 ATTAATCTAGCAATGGTAAATATGACCCGAGTTCCAGGGC 198

RESULT 24
US-08-232-081B-37
; Sequence 37, Application US/08232081B
; Patent No. 5886152
; GENERAL INFORMATION:
; APPLICANT: NAKATANI, TOMOYUKI
; APPLICANT: GOMI, HIDEYUKI
; APPLICANT: WIJENES, JOHN
; APPLICANT: NOGUCHI, HIROSHI
; TITLE OF INVENTION: HUMANIZED B-B10
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,081B
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 20-3484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 354 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..354
US-08-232-081B-37

Query Match 75.3%; Score 38.4; DB 2; Length 354;
Best Local Similarity 87.5%; Pred. No. 1.1e-05;
Matches 42; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 ATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCACGGCC 51
Db 151 ATTAATCTAGCAATGGTAAATATGACCCGAGTTCCAGGGC 198

RESULT 25
US-08-454-899G-19/c
; Sequence 19, Application US/08454899G
; Patent No. 6602503
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/08/454,899G
; CURRENT FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
```





```
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,339A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,128
; FILING DATE: 05-JUN-1995
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-11-1993
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-436-339A-1

Query Match 74.9%; Score 38.2; DB 2; Length 360;
Best Local Similarity 84.3%; Pred. No. 1.3e-05;
Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 AGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 145 AGGATTGATCTCGAGTGGCGACTAAATATGACCCGAGTTCCAGGTC 195

RESULT 29
US-08-950-660-1
; Sequence 1, Application US/08950660
; Patent No. 5932214
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.; Burkly, Linda C.
; TITLE OF INVENTION: Treatment For Inflammatory Bowel Disease
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/950,660
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/456,124
; FILING DATE:
; APPLICATION NUMBER: US 08/373,857
; FILING DATE: 18-JAN-1995
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 08/284,603
; FILING DATE: 11-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00924
; FILING DATE: 02-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,139
; FILING DATE: 12-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis (PLM)
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-031USCP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1...360
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: /note= "pBAG159 insert: HPI/2 heavy
; OTHER INFORMATION: chain variable region; amino acid 1 is Glu (E) but
; OTHER INFORMATION: Gln (Q) may be substituted"
; US-08-950-660-1

Query Match 74.9%; Score 38.2; DB 2; Length 360;
Best Local Similarity 84.3%; Pred. No. 1.3e-05;
Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 AGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 145 AGGATTGATCTCGAGTGGCGACTAAATATGACCCGAGTTCCAGGTC 195

RESULT 30
US-08-454-899G-5
; Sequence 5, Application US/08454899G
; Patent No. 6602503
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/08/454,899G
; CURRENT FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: heavy chain variable region
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(360)
; US-08-454-899G-5

Query Match 74.9%; Score 38.2; DB 4; Length 360;
Best Local Similarity 84.3%; Pred. No. 1.3e-05;
```

Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 145 AGGATTGATCCTCGGAGTGCGGTACTAATAATATGACCCGAAGTTCCAGGTC 195

RESULT 31  
PCT-US93-00030-1  
; Sequence 1, Application PC/TUS9300030  
; GENERAL INFORMATION:  
; APPLICANT: Lobb, Roy R.  
; TITLE OF INVENTION: Treatment for Asthma  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Allegretti & Witcoff, Ltd.  
; STREET: 10 South Wacker Drive, Suite 3000  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: US  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/00030  
; FILING DATE: 19930112  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 92,307-A; D002 CIP PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-715-1000  
; TELEFAX: 312-715-1234  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 360 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1  
; OTHER INFORMATION: /note= "PBAG159 insert: HPI/2 heavy  
; OTHER INFORMATION: chain variable region; amino acid 1 is Glu (E) but Gln (Q)  
; OTHER INFORMATION: may be substituted"  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..360  
; ;  
PCT-US93-00030-1

Query Match 74.9%; Score 38.2; DB 5; Length 360;  
Best Local Similarity 84.3%; Pred. No. 1.3e-05;  
Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 145 AGGATTGATCCTCGGAGTGCGGTACTAATAATATGACCCGAAGTTCCAGGTC 195

RESULT 32  
PCT-US93-00924-1  
; Sequence 1, Application PC/TUS9300924  
; GENERAL INFORMATION:  
; APPLICANT: Lobb, Roy R.  
; TITLE OF INVENTION: Treatment for Inflammatory Bowel Disease  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Allegretti & Witcoff, Ltd.

;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA: PCT/US94/10395  
;; APPLICATION NUMBER: PCT/US94/10395  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA: US 08/122,228  
;; FILING DATE: 15-SEP-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Levine, Leslie M.  
;; REGISTRATION NUMBER: 35,245  
;; REFERENCE/DOCKET NUMBER: D017CIP  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 252-9810  
;; TELEFAX: (617) 252-9617  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 360 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
PCT-US94-10395-1

Query Match 74.9%; Score 38.2; DB 5; Length 360;  
Best Local Similarity 84.3%; Pred. No. 1.3e-05;  
Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 AAGATTGATCTCGAATGCTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 145 AGGATTGATCTCGAGTGGCGATACATAATATGACCCGAGTTCAGGTC 195

RESULT 34  
US-08-822-830B-1  
; Sequence 1, Application US/08822830B  
; Patent No. 5871734  
; GENERAL INFORMATION:  
; APPLICANT: Lobb, Roy R.; Burkly, Linda C.  
; TITLE OF INVENTION: Treatment for Asthma with VLA-4 Blocking  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/822,830B  
; FILING DATE: 03-MAR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/456,193  
; FILING DATE: 31-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/374,331  
; FILING DATE: 18-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,631  
; FILING DATE: 12-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/00030  
; FILING DATE: 12-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/821,768

;; FILING DATE: 13-JAN 1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Myers, Louis (PLM)  
;; REGISTRATION NUMBER: 35,965  
;; REFERENCE/DOCKET NUMBER: BGP-021USCN  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 742-4214  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 363 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..363  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: 1  
;; OTHER INFORMATION: /note= "pBAG159 insert: HP1/2 heavy  
;; OTHER INFORMATION: chain variable region"  
US-08-822-830B-1

Query Match 74.9%; Score 38.2; DB 2; Length 363;  
Best Local Similarity 84.3%; Pred. No. 1.3e-05;  
Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 AAGATTGATCTCGAATGCTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 148 AGGATTGATCTCGAGTGGCGATACATAATATGACCCGAGTTCAGGTC 198

RESULT 35  
US-08-822-830B-12  
; Sequence 12, Application US/08822830B  
; Patent No. 5871734  
; GENERAL INFORMATION:  
; APPLICANT: Lobb, Roy R.; Burkly, Linda C.  
; TITLE OF INVENTION: Treatment for Asthma with VLA-4 Blocking  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/822,830B  
; FILING DATE: 03-MAR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/456,193  
; FILING DATE: 31-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/374,331  
; FILING DATE: 18-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,631  
; FILING DATE: 12-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/00030  
; FILING DATE: 12-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/821,768

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis (PLM)
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-021USCN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 363 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..363
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: /note= "pBAG159 insert: HP1/2 heavy
; OTHER INFORMATION: chain variable region"
US-08-822-830B-12

Query Match 74.9%; Score 38.2; DB 2; Length 363;
Best Local Similarity 84.3%; Pred. No. 1.3e-05;
Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
Db 148 AGGATTGATCTCGGAGTGGCGATACATAATATGACCCGAAGTTCAGGTC 198

RESULT 36
US-09-157-452B-1
; Sequence 1, Application US/09157452B
; Patent No. 6482409
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; TITLE OF INVENTION: TREATMENT FOR INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 10274-004003
; CURRENT APPLICATION NUMBER: US/09/157,452B
; CURRENT FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 08/950,660
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: US 08/373,857
; PRIOR FILING DATE: 1995-01-18
; PRIOR APPLICATION NUMBER: US 08/284,603
; PRIOR FILING DATE: 1994-08-11
; PRIOR APPLICATION NUMBER: PCT/US93/00924
; PRIOR FILING DATE: 1993-02-02
; PRIOR APPLICATION NUMBER: US 07/835,139
; PRIOR FILING DATE: 1992-02-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 363
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(363)
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: /note= "pBAG159 insert: HP1/2 heavy chain variable regions"
US-09-157-452B-15

Query Match 74.9%; Score 38.2; DB 4; Length 363;
Best Local Similarity 84.3%; Pred. No. 1.3e-05;
Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
Db 148 AGGATTGATCTCGGAGTGGCGATACATAATATGACCCGAAGTTCAGGTC 198

RESULT 37
US-09-157-452B-15
; Sequence 15, Application US/09157452B
; Patent No. 6482409
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; TITLE OF INVENTION: TREATMENT FOR INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 10274-004003
; CURRENT APPLICATION NUMBER: US/09/157,452B
; CURRENT FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 08/950,660
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: US 08/373,857
; PRIOR FILING DATE: 1995-01-18
; PRIOR APPLICATION NUMBER: US 08/284,603
; PRIOR FILING DATE: 1994-08-11
; PRIOR APPLICATION NUMBER: PCT/US93/00924
; PRIOR FILING DATE: 1993-02-02
; PRIOR APPLICATION NUMBER: US 07/835,139
; PRIOR FILING DATE: 1992-02-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 363
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(363)
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: /note= "pBAG159 insert: HP1/2 heavy chain variable regions"
US-09-157-452B-15

Query Match 74.9%; Score 38.2; DB 4; Length 363;
Best Local Similarity 84.3%; Pred. No. 1.3e-05;
Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
Db 148 AGGATTGATCTCGGAGTGGCGATACATAATATGACCCGAAGTTCAGGTC 198

RESULT 38
US-08-454-899G-46
; Sequence 46, Application US/08454899G
; Patent No. 6602503
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/08/454,899G
; CURRENT FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(363)
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: /note= "pBAG159 insert: HP1/2 heavy chain variable regions;
; OTHER INFORMATION: amino acid 1 is Glu (E) but Gln (Q) may be substituted as shown i
; OTHER INFORMATION: SEQ ID NO:15 and 16"
US-09-157-452B-1

Query Match 74.9%; Score 38.2; DB 4; Length 363;
Best Local Similarity 84.3%; Pred. No. 1.3e-05;
Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 1 AAGATTGATCTCGGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
Db 148 AGGATTGATCTCGGAGTGGCGATACATAATATGACCCGAAGTTCAGGTC 198

RESULT 37
US-09-157-452B-15
; Sequence 15, Application US/09157452B
; Patent No. 6482409
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; TITLE OF INVENTION: TREATMENT FOR INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 10274-004003
; CURRENT APPLICATION NUMBER: US/09/157,452B
; CURRENT FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 08/950,660
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: US 08/373,857
; PRIOR FILING DATE: 1995-01-18
; PRIOR APPLICATION NUMBER: US 08/284,603
; PRIOR FILING DATE: 1994-08-11
; PRIOR APPLICATION NUMBER: PCT/US93/00924
; PRIOR FILING DATE: 1993-02-02
; PRIOR APPLICATION NUMBER: US 07/835,139
; PRIOR FILING DATE: 1992-02-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 363
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(363)
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: /note= "pBAG159 insert: HP1/2 heavy chain variable regions"
US-09-157-452B-15

Query Match 74.9%; Score 38.2; DB 4; Length 363;
Best Local Similarity 84.3%; Pred. No. 1.3e-05;
Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
Db 148 AGGATTGATCTCGGAGTGGCGATACATAATATGACCCGAAGTTCAGGTC 198

RESULT 38
US-08-454-899G-46
; Sequence 46, Application US/08454899G
; Patent No. 6602503
; GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Carr, Frank J.
; APPLICANT: Tempest, Philip R.
; TITLE OF INVENTION: RECOMBINANT ANTI-VLA4 ANTIBODY MOLECULES
; FILE REFERENCE: 10274-007001
; CURRENT APPLICATION NUMBER: US/08/454,899G
; CURRENT FILING DATE: 1995-05-31
; PRIOR APPLICATION NUMBER: US 08/004,798
; PRIOR FILING DATE: 1993-01-12
; PRIOR APPLICATION NUMBER: PCT/US94/00266
; PRIOR FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```



**Qy** 1 AAGATTGATCTGCGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
**Dd** 205 AGSATTGATCTGCGAGTGCGGATACTAAATATGACCCGAAGTTCCAGTGC 255

Search completed: June 4, 2005, 15:59:10  
Job time : 101.455 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 14:31:52 ; Search time 403.364 Seconds  
(without alignments)  
777.260 Million cell updates/sec

Title: US-10-089-452-25

Perfect score: 51

Sequence: 1 aagattgatctgcgaatgg.....atgaccgatattccaggcc 51

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18:	/cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	84.3	51	17	US-10-344-779-17
2	43	84.3	51	18	US-10-738-809-17
3	43	84.3	279	14	US-10-179-851-6
4	43	84.3	357	18	US-10-481-212-2
5	43	84.3	399	17	US-10-344-779-3
6	43	84.3	399	18	US-10-738-809-3
7	43	84.3	412	14	US-10-283-349-30
8	43	84.3	423	19	US-10-687-035-58
9	41.4	81.2	132	10	US-09-749-873-95
10	41.4	81.2	330	14	US-10-179-851-1
11	41.4	81.2	351	14	US-10-059-261-220
					Sequence 17, Appl
					Sequence 17, Appl
					Sequence 6, Appl
					Sequence 2, Appl
					Sequence 3, Appl
					Sequence 3, Appl
					Sequence 30, Appl
					Sequence 58, Appl
					Sequence 95, Appl
					Sequence 1, Appl
					Sequence 220, App

12	41.4	81.2	351	18	US-10-627-649-220	Sequence 220, App
13	41.4	81.2	351	19	US-10-855-013-25	Sequence 25, Appl
14	41.4	81.2	360	15	US-10-056-794-14	Sequence 14, Appl
15	41.4	81.2	360	19	US-10-787-067-14	Sequence 14, Appl
16	41.4	81.2	360	19	US-10-631-660-14	Sequence 14, Appl
17	41.4	81.2	363	19	US-10-879-994-33	Sequence 33, Appl
18	41.4	81.2	409	10	US-09-749-873-28	Sequence 28, Appl
19	41.4	81.2	409	10	US-09-749-873-98	Sequence 98, Appl
20	41.4	81.2	454	19	US-10-763-424-66	Sequence 66, Appl
21	41.4	81.2	470	19	US-10-763-424-53	Sequence 53, Appl
22	41.4	81.2	822	10	US-09-749-873-108	Sequence 108, App
23	39.8	78.0	357	8	US-08-996-140-19	Sequence 19, Appl
24	39.8	78.0	414	8	US-08-996-140-21	Sequence 21, Appl
25	39.8	78.0	783	14	US-10-013-173-36	Sequence 36, Appl
26	39.8	78.0	783	15	US-10-150-762-36	Sequence 36, Appl
27	39.8	78.0	783	15	US-10-244-821-36	Sequence 36, Appl
28	39.8	78.0	5711	9	US-09-897-006-8	Sequence 8, Appl
29	39.8	78.0	5711	10	US-09-897-511A-8	Sequence 8, Appl
30	39.8	78.0	5711	17	US-10-397-079-8	Sequence 8, Appl
31	39.8	78.0	5711	18	US-10-759-315-8	Sequence 8, Appl
32	39.8	78.0	5711	19	US-10-947-881-8	Sequence 8, Appl
33	39.8	78.0	5711	20	US-11-018-895-8	Sequence 8, Appl
34	39.8	78.0	6255	9	US-09-897-006-13	Sequence 13, Appl
35	39.8	78.0	6255	10	US-09-897-511A-13	Sequence 13, Appl
36	39.8	78.0	6255	17	US-10-397-079-13	Sequence 13, Appl
37	39.8	78.0	6255	18	US-10-759-315-13	Sequence 13, Appl
38	39.8	78.0	6255	19	US-10-947-881-13	Sequence 13, Appl
39	39.8	78.0	6255	20	US-11-018-895-13	Sequence 13, Appl
40	39.4	77.3	771	17	US-10-239-656-66	Sequence 66, Appl
41	38.2	74.9	71	16	US-10-428-662-19	Sequence 19, Appl
42	38.2	74.9	84	18	US-10-487-525-6	Sequence 6, Appl
43	38.2	74.9	348	18	US-10-482-630-49	Sequence 49, Appl
44	38.2	74.9	360	16	US-10-428-662-5	Sequence 5, Appl
45	38.2	74.9	360	19	US-10-875-282-1	Sequence 1, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-344-779-17  
; Sequence 17, Application US/10344779  
; Publication No. US20030211106A1  
; GENERAL INFORMATION:  
; APPLICANT: SWEET, RAYMOND W.  
; APPLICANT: TORNETTA, MARK A.  
; APPLICANT: TRUNEH, ALEMSEGED  
; APPLICANT: WATTAM, TREVOR A.  
; TITLE OF INVENTION: ANTI-RANK LIGAND MONOCLONAL ANTIBODIES  
; FILE REFERENCE: GP50033  
; CURRENT APPLICATION NUMBER: US/10/344,779  
; CURRENT FILING DATE: 2003-02-17  
; PRIOR APPLICATION NUMBER: PCT/US01/26161  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,524  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 60/230,639  
; PRIOR FILING DATE: 2000-09-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 51  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-344-779-17

Query Match 84.3%; Score 43; DB 17; Length 51;  
Best Local Similarity 90.2%; Pred. No. 6e-06;  
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAGATTGATCTGCGAATGGTAAACTAAATATGACCGGATATTCAGGCC 51  
|||||

Db 1 AGGATTGATCTCGGAATGGTAATACTAAATATGACCCGGAAGTTCCAGGCC 51

## RESULT 2

US-10-738-809-17  
; Sequence 17, Application US/10738809  
; Publication No. US20040171117A1  
; GENERAL INFORMATION:  
; APPLICANT: SWEET, RAYMOND W.  
; APPLICANT: TORNETTA, MARK A.  
; APPLICANT: TRUNEH, ALEMSEGED  
; APPLICANT: WAITAM, TREVOR A.  
; TITLE OF INVENTION: ANTI-RANK LIGAND MONOCLONAL ANTIBODIES  
; TITLE OF INVENTION: USEFUL IN TREATMENT OF RANK LIGAND MEDIATED DISORDERS  
; FILE REFERENCE: GP50033  
; CURRENT APPLICATION NUMBER: US/10/738,809  
; CURRENT FILING DATE: 2003-12-17  
; PRIOR APPLICATION NUMBER: US/10/344,779  
; PRIOR FILING DATE: 2003-02-19  
; PRIOR APPLICATION NUMBER: PCT/US01/26161  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,524  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 60/230,639  
; PRIOR FILING DATE: 2000-09-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 51  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-738-809-17

Query Match 84.3%; Score 43; DB 18; Length 51;  
Best Local Similarity 90.2%; Pred. No. 6e-06; Mismatches 5; Indels 0; Gaps 0;  
Matches 46; Conservative 0;

Qy 1 AGAGTTGATCTCGGAATGGTAATACTAAATATGACCCGGAATATTCAGGCC 51  
Db 1 AGGATTGATCTCGGAATGGTAATACTAAATATGACCCGGAAGTTCCAGGCC 51

## RESULT 3

US-10-179-851-6  
; Sequence 6, Application US/10179851  
; Publication No. US20030017518A1  
; GENERAL INFORMATION:  
; APPLICANT: Lam et al.  
; APPLICANT: Lam, Joseph S  
; APPLICANT: Zhao, Xin S  
; TITLE OF INVENTION: NON-RADIOACTIVE ASSAY OF LIPOPOLYSACCHARIDE KINASES  
; FILE REFERENCE: BEP 3004  
; CURRENT APPLICATION NUMBER: US/10/179,851  
; CURRENT FILING DATE: 2002-08-20  
; PRIOR APPLICATION NUMBER: US 60/300,420  
; PRIOR FILING DATE: 2001-06-26  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 6  
; LENGTH: 279  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(279)  
; OTHER INFORMATION:  
US-10-179-851-6

Query Match 84.3%; Score 43; DB 14; Length 279;  
Best Local Similarity 90.2%; Pred. No. 8.7e-06; Mismatches 5; Indels 0; Gaps 0;  
Matches 46; Conservative 0;

Qy 1 AGAGTTGATCTCGGAATGGTAATACTAAATATGACCCGGAATATTCAGGCC 51

Db 133 AGGATTGATCTCGGAATGGTAATACTAAATATGACCCGGAAGTTCCAGGCC 183

## RESULT 4

US-10-481-212-2  
; Sequence 2, Application US/10481212  
; Publication No. US20040214761A1  
; GENERAL INFORMATION:  
; APPLICANT: PacMab Pty Ltd  
; TITLE OF INVENTION: Method for treating multiple myeloma  
; FILE REFERENCE: 500564  
; CURRENT APPLICATION NUMBER: US/10/481,212  
; CURRENT FILING DATE: 2004-06-04  
; PRIOR APPLICATION NUMBER: AU PR 6179  
; PRIOR FILING DATE: 2001-07-06  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 2  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-481-212-2

Query Match 84.3%; Score 43; DB 18; Length 357;  
Best Local Similarity 90.2%; Pred. No. 9.1e-06; Mismatches 5; Indels 0; Gaps 0;  
Matches 46; Conservative 0;

Qy 1 AGAGTTGATCTCGGAATGGTAATACTAAATATGACCCGGAATATTCAGGCC 51  
Db 148 AGGATTGATCTCGGAATGGTAATACTAAATATGACCCGGAAGTTCCAGGCC 198

## RESULT 5

US-10-344-779-3  
; Sequence 3, Application US/10344779  
; Publication No. US20030211106A1  
; GENERAL INFORMATION:  
; APPLICANT: SWEET, RAYMOND W.  
; APPLICANT: TORNETTA, MARK A.  
; APPLICANT: TRUNEH, ALEMSEGED  
; APPLICANT: WAITAM, TREVOR A.  
; TITLE OF INVENTION: ANTI-RANK LIGAND MONOCLONAL ANTIBODIES  
; TITLE OF INVENTION: USEFUL IN TREATMENT OF RANK LIGAND MEDIATED DISORDERS  
; FILE REFERENCE: GP50033  
; CURRENT APPLICATION NUMBER: US/10/344,779  
; CURRENT FILING DATE: 2003-02-17  
; PRIOR APPLICATION NUMBER: PCT/US01/26161  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,524  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 60/230,639  
; PRIOR FILING DATE: 2000-09-07  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 399  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-344-779-3

Query Match 84.3%; Score 43; DB 17; Length 399;  
Best Local Similarity 90.2%; Pred. No. 9.4e-06; Mismatches 5; Indels 0; Gaps 0;  
Matches 46; Conservative 0;

Qy 1 AGAGTTGATCTCGGAATGGTAATACTAAATATGACCCGGAATATTCAGGCC 51  
Db 148 AGGATTGATCTCGGAATGGTAATACTAAATATGACCCGGAAGTTCCAGGCC 198

## RESULT 6

US-10-738-809-3  
; Sequence 3, Application US/10738809



```

; NAME: Lawrence, III, Stanton T
; REGISTRATION NUMBER: 25,736
; REFERENCE/DOCKET NUMBER: 7005-115-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; SEQUENCE DESCRIPTION: SEQ ID NO: 30:
; US-10-283-349--30
;
Query Match      84.3%; Score 43; DB 14; Length 412;
Best Local Similarity 90.2%; Pred. No. 9.4e-06;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      1  AGAATTGATCCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51
      |||||
Db      205 AGGAATTGATCCTCGGAATGGTAACTAAATCTGACCCGAAGTTCCAGGCC 255

```

US-10-687-035-58  
; Sequence 58, Application US/10687035  
; Publication No. US20050064518A1

APPLICANT: Albone, Earl F.  
APPLICANT: Soltis, Daniel A.  
TITLE OF INVENTION: ANTIBODIES THAT BIND CELL-ASSOCIATED

```

: CURRENT APPLICATION NUMBER: US/10/687,035
: CURRENT FILING DATE: 2003-10-15
: PRIOR APPLICATION NUMBER: 60/485,986
: PRIOR FILING DATE: 2003-07-10
: PRIOR APPLICATION NUMBER: 60/418,828
: PRIOR FILING DATE: 2003-10-12
: NUMBER OF SEQ ID NOS: 71
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 58
: LENGTH: 423
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: 16H9 heavy chain polypeptide variable region (16H9H)

```

```

US-10-687-035-58

Query Match      84.3%  Score 43;  DB 19;  Length 423;
Best Local Similarity 90.2%;  Pred. No. 9.5e-06;
Matches 46;  Conservative 0;  Mismatches 5;  Indels 0;  Gaps 0;

QY      1  AAGATTGATCCTCGGAATGTTAAACTAAATATGACCCGATATTCAGGGC 51
      |||||
Db       205 AGGATTGATCCTCGGAATGTTAACTAAATATGACCCGAAGTTCAGGGC 255

RESULT 9
US-09-749-873-95/c
; Sequence 95, Application US/09749873
; Publication No. US20030023045A1
; GENERAL INFORMATION:
; APPLICANT: OHTOMO, Toshihiko
; SATO, Koh
; TSUCHIYA, Masayuki
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
; MEDULLOBLASTOMA CELLS
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
;

```

RESULT 9  
US-09-749-873-95/c  
; Sequence 95, Application US/09749873  
; Publication No. US20030023045A1  
; GENERAL INFORMATION:  
; APPLICANT: OHTOMO, Toshihiko  
; SATO, Koh  
; TSUCHIYA, Masayuki  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; MEDULLOBLASTOMA CELLS  
; NUMBER OF SEQUENCES: 132  
; CORRESPONDENCE ADDRESS:  
;

```
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/749,873
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/646,265
; FILING DATE: 1996-09-09
; APPLICATION NUMBER: JP 5-291078
; FILING DATE: 19-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 95:
US-09-749-873-95

Query Match      81.2%; Score 41.4; DB 10; Length 132;
Best Local Similarity 88.2%; Pred. No. 2.8e-05;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
Db 132 AGGATTGATCCTCGGATGTAATACTAAATATGACCCGAGTTCCAGGCC 82

RESULT 10
US-10-179-851-1
; Sequence 1, Application US/10179851
; Publication No. US20030017518A1
; GENERAL INFORMATION:
; APPLICANT: Lam et al.
; APPLICANT: Lam, Joseph S
; APPLICANT: Zhao, Xin S
; TITLE OF INVENTION: NON-RADIOACTIVE ASSAY OF LIPOPOLYSACCHARIDE KINASES
; FILE REFERENCE: BEP 3004
; CURRENT APPLICATION NUMBER: US/10/179,851
; CURRENT FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: US 60/300,420
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(330)
; OTHER INFORMATION:
US-10-179-851-1

Query Match      81.2%; Score 41.4; DB 14; Length 330;
Best Local Similarity 88.2%; Pred. No. 3.5e-05;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
Db 148 AGGATTGATCCTCGAATGTTAAACTAAATATGACCCGAGTTCCAGGCC 198

RESULT 12
US-10-627-649-220
; Sequence 220, Application US/10627649
; Publication No. US20040265300A1
; GENERAL INFORMATION:
; APPLICANT: EDELMAN, LENA
; APPLICANT: JACOTOT, ETIENNE DANIEL FRANCOIS
; APPLICANT: BRIAND, JEAN-PAUL
; TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
; TITLE OF INVENTION: SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
; TITLE OF INVENTION: FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
; FILE REFERENCE: (PTPC)
; FILE REFERENCE: 02356-0083
; CURRENT APPLICATION NUMBER: US/10/627,649
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: 60/265,594
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 220
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic VH
; OTHER INFORMATION: nucleotide sequence
US-10-627-649-220

Query Match      81.2%; Score 41.4; DB 14; Length 351;
Best Local Similarity 88.2%; Pred. No. 3.5e-05;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
Db 148 AGGATTGATCCTCGAATGTTAAACTAAATATGACCCGAGTTCCAGGCC 198

RESULT 11
US-10-059-261-220
; Sequence 220, Application US/10059261
; Publication No. US20030077826A1
; GENERAL INFORMATION:
; APPLICANT: EDELMAN, LENA
; APPLICANT: JACOTOT, ETIENNE DANIEL FRANCOIS
; APPLICANT: BRIAND, JEAN-PAUL
; TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
; TITLE OF INVENTION: SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
; TITLE OF INVENTION: FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
; TITLE OF INVENTION: (PTPC)
; FILE REFERENCE: 03495,0216
; CURRENT APPLICATION NUMBER: US/10/059,261
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/265,594
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 220
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic VH
; OTHER INFORMATION: nucleotide sequence
US-10-059-261-220

Query Match      81.2%; Score 41.4; DB 14; Length 351;
Best Local Similarity 88.2%; Pred. No. 3.5e-05;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51
Db 148 AGGATTGATCCTCGAATGTTAAACTAAATATGACCCGAGTTCCAGGCC 198

RESULT 12
US-10-627-649-220
; Sequence 220, Application US/10627649
; Publication No. US20040265300A1
; GENERAL INFORMATION:
; APPLICANT: EDELMAN, LENA
; APPLICANT: JACOTOT, ETIENNE DANIEL FRANCOIS
; APPLICANT: BRIAND, JEAN-PAUL
; TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
; TITLE OF INVENTION: SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
; TITLE OF INVENTION: FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
; FILE REFERENCE: (PTPC)
; FILE REFERENCE: 02356-0083
; CURRENT APPLICATION NUMBER: US/10/627,649
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: 60/265,594
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 220
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic VH
; OTHER INFORMATION: nucleotide sequence
US-10-627-649-220
```



```

RESULT 17
US-10-879-994-33
; Sequence 33, Application US/10879994
; Publication No. US20050032175A1
; GENERAL INFORMATION:
; APPLICANT: Stahl, Neil
; APPLICANT: Yancopoulos, George D.
; APPLICANT: Karow, Margaret
; APPLICANT: Smith, Eric
; TITLE OF INVENTION: HIGH AFFINITY FUSION PROTEINS AND THERAPEUTIC AND DIAGNOSTIC METHODS
; TITLE OF INVENTION: US
; FILE REFERENCE: REG 203E2
; CURRENT APPLICATION NUMBER: US/10/879,994
; CURRENT FILING DATE: 2004-06-29
; PRIOR APPLICATION NUMBER: 10/610,452
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 363
; TYPE: DNA
; ORGANISM: mus musculus
US-10-879-994-33

Query Match      81.2%; Score 41.4; DB 19; Length 363;
Best Local Similarity 88.2%; Pred. No. 3.5e-05;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      1  AGATTGATCTCGGAATGTTAAATAATATGACCCGATATTCAGGCC 51
      |||
Db      148 AGGATTGATCTCGGAATGTTAATACTAAATATGACCCGAAAGTTCAGGCC 198

RESULT 18
US-09-749-873-28
; Sequence 28, Application US/09749873
; Publication No. US20030023045A1
; GENERAL INFORMATION:
; APPLICANT: OHTOMO, Toshihiko
; TATO, Koh
; TSUCHIYA, Masayuki
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
; MEDULLOBLASTOMA CELLS
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/749,873
; FILING DATE: 29-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/646,265
; FILING DATE: 1996-09-09
; APPLICATION NUMBER: JP 5-291078
; FILING DATE: 19-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/184
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399

```

```

; LOCATION: 1..408
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 58..408
; SEQUENCE DESCRIPTION: SEQ ID NO: 98:
US-09-749-873--98

Query Match      81.2%; Score 41.4; DB 10; Length 409;
Best Local Similarity 88.2%; Pred. No. 3.6e-05;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGGAATGCTAAAACTAAATATGACCCGATATTCAGGCC 51
Db 205 AGGATTGATCCTCGGAATGCTAAATATGACCCGAAAGTTCAGGCC 255

RESULT 20
US-10-763-424-66
; Sequence 66, Application US/10763424
; Publication No. US20050069541A1
; GENERAL INFORMATION:
; APPLICANT: Karlik, Stephen J.
; APPLICANT: Pleiss, Michael A.
; APPLICANT: Konradi, Andrei W.
; APPLICANT: Grant, Francine S.
; APPLICANT: Semko, Christopher M.
; APPLICANT: Dreszen, Darren B.
; APPLICANT: Messersmith, Elizabeth
; APPLICANT: Freedman, Stephen
; APPLICANT: Yednock, Ted
; TITLE OF INVENTION: Composition for and Treatment of Demyelinating Diseases
; FILE REFERENCE: 034008-061
; CURRENT APPLICATION NUMBER: US/10/763,424
; CURRENT FILING DATE: 2004-01-26
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 454
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16)...(441)
US-10-763-424-66

Query Match      81.2%; Score 41.4; DB 19; Length 454;
Best Local Similarity 88.2%; Pred. No. 3.7e-05;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGGAATGCTAAAACTAAATATGACCCGATATTCAGGCC 51
Db 220 AGGATTGATCCTCGGAATGCTAAATATGACCCGAAAGTTCAGGCC 270

RESULT 21
US-10-763-424-53
; Sequence 53, Application US/10763424
; Publication No. US20050069541A1
; GENERAL INFORMATION:
; APPLICANT: Karlik, Stephen J.
; APPLICANT: Pleiss, Michael A.
; APPLICANT: Konradi, Andrei W.
; APPLICANT: Grant, Francine S.
; APPLICANT: Semko, Christopher M.
; APPLICANT: Dreszen, Darren B.
; APPLICANT: Messersmith, Elizabeth
; APPLICANT: Freedman, Stephen
; APPLICANT: Yednock, Ted
; TITLE OF INVENTION: Composition for and Treatment of Demyelinating Diseases
; FILE REFERENCE: 034008-061
; CURRENT APPLICATION NUMBER: US/10/763,424

```

;; CURRENT FILING DATE: 2004-01-26  
;; NUMBER OF SEQ ID NOS: 71  
;; SOFTWARE: FastSEQ for Windows Version 4.0  
;; SEQ ID NO 53  
;; LENGTH: 470  
;; TYPE: DNA  
;; ORGANISM: Mouse  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (1)...(432)  
US-10-763-424-53

Query Match 81.2%; Score 41.4; DB 19; Length 470;  
Best Local Similarity 88.2%; Pred. No. 3.7e-05;  
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 AAGATTGATCTCGGAATGTTAAACTAAATATGACCCGATATTCACGGCC 51  
Db 205 AGGATTGATCTCGGAATGTTATCTACTAAATATGACCCGAGTTCACGGCC 255

RESULT 22  
US-09-749-873-108  
; Sequence 108 Application US/09749873  
; Publication No. US20030023045A1  
; GENERAL INFORMATION:  
; APPLICANT: OHTOMO, Toshihiko  
; SATO, Koh  
; TSUCHIYA, Masayuki  
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
; MEDULLOBLASTOMA CELLS  
; NUMBER OF SEQUENCES: 132  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/749,873  
; FILING DATE: 29-Dec-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/646,265  
; FILING DATE: 1996-09-09  
; APPLICATION NUMBER: JP 5-291078  
; FILING DATE: 19-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WEGNER, Harold C.  
; REGISTRATION NUMBER: 25,258  
; REFERENCE/DOCKET NUMBER: 53466/184  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 108:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 822 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..807  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 1..807

;; SEQUENCE DESCRIPTION: SEQ ID NO: 108:  
US-09-749-873-108

Query Match 81.2%; Score 41.4; DB 10; Length 822;  
Best Local Similarity 88.2%; Pred. No. 4.2e-05;  
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 AAGATTGATCTCGGAATGTTAAACTAAATATGACCCGATATTCACGGCC 51  
Db 214 AGGATTGATCTCGGAATGTTATCTACTAAATATGACCCGAGTTCACGGCC 264

RESULT 23  
US-08-996-140-19  
; Sequence 19 Application US/08996140  
; Publication No. US20030190318A1  
; GENERAL INFORMATION:  
; APPLICANT: TORIGOE, Kakuji  
; APPLICANT: USHIO, Shimpel  
; APPLICANT: KURIKATA, Toshio  
; APPLICANT: KURIMOTO, Masaashi  
; TITLE OF INVENTION: INTERLEUKIN-18 RECEPTOR PROTEINS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,140  
; FILING DATE: 22-DEC-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 356,426/1996  
; FILING DATE: 26-DEC-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 52,526/1997  
; FILING DATE: 21-FEB-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 163,490/1997  
; FILING DATE: 6-JUN-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 215,490/1997  
; FILING DATE: 28-JUL-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROWDY, Roger L.  
; REGISTRATION NUMBER: 25,618  
; REFERENCE/DOCKET NUMBER: TORIGOE-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 357 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 1..357  
; IDENTIFICATION METHOD: E  
US-08-996-140-19

Query Match 78.0%; Score 39.8; DB 8; Length 357;  
Best Local Similarity 86.3%; Pred. No. 0.00014;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTTCCAGGCC 51  
Db 148 AGGATTGATCTCGAATGGTAACTAAATATGACCCGATATTTCCAGGC 198

RESULT 24  
US-08-996-140-21  
; Sequence 21, Application US/08996140  
; Publication No. US20030190318A1  
; GENERAL INFORMATION:  
; APPLICANT: TORIGOE, Kakuji  
; APPLICANT: USHIO, Shimpel  
; APPLICANT: KONIKATA, Toshio  
; APPLICANT: KURIMOTO, Masashi  
; TITLE OF INVENTION: INTERLEUKIN-18 RECEPTOR PROTEINS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,140  
; FILING DATE: 22-DEC-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 356,426/1996  
; FILING DATE: 26-DEC-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 52,526/1997  
; FILING DATE: 21-FEB-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 163,490/1997  
; FILING DATE: 6-JUN-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 215,490/1997  
; FILING DATE: 28-JUL-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROWDY, Roger L.  
; REGISTRATION NUMBER: 25,618  
; REFERENCE/DOCKET NUMBER: TORIGOE-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 414 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: sig peptide  
; LOCATION: 1..57  
; IDENTIFICATION METHOD: E  
; NAME/KEY: mat peptide  
; LOCATION: 58..414  
; IDENTIFICATION METHOD: E  
US-08-996-140-21

Query Match 78.0%; Score 39.8; DB 8; Length 414;  
Best Local Similarity 86.3%; Pred. No. 0.00014;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTTCCAGGCC 51

Db 205 AGGATTGATCTCGAATGGTAACTAAATATGACCCGATATTTCCAGGC 255

RESULT 25  
US-10-013-173-36  
; Sequence 36, Application US/10013173  
; Publication No. US20030095977A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen C.  
; APPLICANT: Graves, Scott Stoll  
; APPLICANT: Schultz, Joanne Elaine  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James A.  
; APPLICANT: Reno, Jorh M.  
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
; FILE REFERENCE: 690022.547C1  
; CURRENT APPLICATION NUMBER: US/10/013,173  
; CURRENT FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 783  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: T84.66 single chain antibody-genomic Streptavidin  
; OTHER INFORMATION: fusion construct  
US-10-013-173-36

Query Match 78.0%; Score 39.8; DB 14; Length 783;  
Best Local Similarity 86.3%; Pred. No. 0.00016;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTTCCAGGCC 51  
Db 148 AGGATTGATCTCGAATGGTAACTAAATATGACCCGATATTTCCAGGC 198

RESULT 26  
US-10-150-762-36  
; Sequence 36, Application US/10150762  
; Publication No. US20030103948A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen C.  
; APPLICANT: Graves, Scott S.  
; APPLICANT: Schultz, Joanne E.  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James A.  
; APPLICANT: Reno, Jorh M.  
; APPLICANT: Dearstyn, Erica A.  
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
; FILE REFERENCE: 690022.547C2  
; CURRENT APPLICATION NUMBER: US/10/150,762  
; CURRENT FILING DATE: 2002-05-17  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 783  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: T84.66 single chain antibody-genomic Streptavidin  
; OTHER INFORMATION: fusion construct  
US-10-150-762-36

Query Match 78.0%; Score 39.8; DB 15; Length 783;  
Best Local Similarity 86.3%; Pred. No. 0.00016;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGAATGGTAAACTAAATATGACCCGATATTTCCAGGCC 51

```

Db      148 AGGATTGATCTCCGAACTGGTAATAGTAAATATGTGCCGAAGTTCCAGGGC 198

RESULT 27
US-10-244-821-36
; Sequence 36, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearatyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: T84.66 single chain antibody-genomic Streptavi
US-10-244-821-36

```

Query Match 78.0%; Score 39.8; DB 15; Length 783;  
Best Local Similarity 86.3%; Pred. No. 0.00016;  
Matches 44; Conservative 0; Mismatches 7; Indels 0;

QY 1 AAGATTGATCTCGGAATGGTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 148 AGGATTGATCTCGGAATGGTAAATAGTAAATATGTCGGAATTCAGGCC 198

```

RESULT 28
US-09-897-006-8
; Sequence 8, Application US/09897006
; Patent No. US20020106729A1
; GENERAL INFORMATION:
; APPLICANT: Block, Gregory
; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/09/897,006
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 5711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-006-8

```

Query Match 78.0%; Score 39.8; DB 9; Length 5711;  
Best Local Similarity 86.3%; Pred. No. 0.00025;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0

QY 1 AAGATTGATCTGGCAATGTGAAAACTAAATATGACCCGATATTCAGGCC 51  
DB 3892 AGGATTGATCTGGCAATGGGAATCTGATATGATGACCCGAAGTTTCAGGCC 3942

RESULT 29  
US-09-897-511A-8

```

; Sequence 8, Application US/09897511A
; Publication No. US20030092882A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
;
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
;
; FILE REFERENCE: GALA-06416
;
; CURRENT APPLICATION NUMBER: US/09/897,511A
; CURRENT FILING DATE: 2001-06-29
;
; PRIOR APPLICATION NUMBER: 60/215,925
; PRIOR FILING DATE: 2000-07-03
;
; NUMBER OF SEQ ID NOS: 36
;
; SOFTWARE: PatentIn version 3.0
;
; SEQ ID NO 8
;
; LENGTH: 5711
;
; TYPE: DNA
;
; ORGANISM: Artificial Sequence
;
; FEATURE:
;
;   OTHER INFORMATION: Synthetic
;
; US-09-897-511A-8

```

```

RESULT 30
US-10-397-079-8
; Sequence 8, Application US/10397079
; Publication No. US20030224415A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/10/397,079
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US/09/897,511A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 5711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-397-079-8

```

RESULT 31  
US-10-759-315-8  
; Sequence 8, Application US/10759315  
; Publication NO. US20040235173A1  
; GENERAL INFORMATION:  
; APPLICANT: Bleck, Gregory T.  
; APPLICANT: Bremel, Robert D.



```
; APPLICANT: Miller, Linda U.
; TITLE OF INVENTION: Production of Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-08484
; CURRENT APPLICATION NUMBER: US/10/759,315
; PRIOR FILING DATE: 2004-01-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 5711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-759-315-8

Query Match      78.0%; Score 39.8; DB 18; Length 5711;
Best Local Similarity 86.3%; Pred. No. 0.00025;
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGAATGCTAAAACTAAATATGACCCGATATTCAGGCC 51
Db 3892 AGGATTGATCCTCGAATGCGAATGCTAAATGACCCGATGCTGAATATGACCCGATTCAGGCC 3942

RESULT 32
US-10-947-881-8
; Sequence 8, Application US/10947881
; Publication No. US2005006762A1
; GENERAL INFORMATION:
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/10/947,881
; CURRENT FILING DATE: 2004-09-23
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/897,006
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 5711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-947-881-8

Query Match      78.0%; Score 39.8; DB 19; Length 5711;
Best Local Similarity 86.3%; Pred. No. 0.00025;
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGAATGCTAAAACTAAATATGACCCGATATTCAGGCC 51
Db 3892 AGGATTGATCCTCGAATGCGAATGCTAAATGACCCGATGCTGAATATGACCCGATTCAGGCC 3942

RESULT 33
US-11-018-895-8
; Sequence 8, Application US/11018895
; Publication No. US20050100952A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/11/018,895
; CURRENT FILING DATE: 2004-12-21
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US/09/897,511
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
```

```
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 5711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-11-018-895-8

Query Match      78.0%; Score 39.8; DB 20; Length 5711;
Best Local Similarity 86.3%; Pred. No. 0.00025;
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGAATGCTAAAACTAAATATGACCCGATATTCAGGCC 51
Db 3892 AGGATTGATCCTCGAATGCGAATGCTAAATGACCCGATGCTGAATATGACCCGATTCAGGCC 3942

RESULT 34
US-09-897-006-13
; Sequence 13, Application US/09897006
; Patent No. US20020106729A1
; GENERAL INFORMATION:
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/09/897,006
; CURRENT FILING DATE: 2001-06-29
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 6255
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-006-13

Query Match      78.0%; Score 39.8; DB 9; Length 6255;
Best Local Similarity 86.3%; Pred. No. 0.00025;
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGAATGCTAAAACTAAATATGACCCGATATTCAGGCC 51
Db 5021 AGGATTGATCCTCGAATGCGAATGCTGAATATGACCCGATTCAGGCC 5071

RESULT 35
US-09-897-511A-13
; Sequence 13, Application US/09897511A
; Publication No. US20030092882A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/09/897,511A
; CURRENT FILING DATE: 2001-06-29
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 6255
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
```

US-09-897-511A-13

Query Match 78.0%; Score 39.8; DB 10; Length 6255;  
Best Local Similarity 86.3%; Pred. No. 0.00025;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 5021 AGGATTGATCCTCGAATGGAATACTGTAATATGACCCGAAAGTTCCAGGCC 5071

RESULT 36  
US-10-397-079-13  
; Sequence 13, Application US/10397079  
; Publication No. US20030224415A1  
; GENERAL INFORMATION:  
; APPLICANT: Bremel, Robert  
; APPLICANT: Miller, Linda  
; APPLICANT: Bleck, Gregory  
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors  
; FILE REFERENCE: GALA-06416  
; CURRENT APPLICATION NUMBER: US/10/397,079  
; PRIOR FILING DATE: 2003-03-26  
; PRIOR APPLICATION NUMBER: US/09/897,511A  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 60/215,925  
; PRIOR FILING DATE: 2000-07-03  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13  
; LENGTH: 6255  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-397-079-13

Query Match 78.0%; Score 39.8; DB 17; Length 6255;  
Best Local Similarity 86.3%; Pred. No. 0.00025;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 5021 AGGATTGATCCTCGAATGGAATACTGTAATATGACCCGAAAGTTCCAGGCC 5071

RESULT 37  
US-10-759-315-13  
; Sequence 13, Application US/10759315  
; Publication No. US20040235173A1  
; GENERAL INFORMATION:  
; APPLICANT: Bleck, Gregory T.  
; APPLICANT: Bremel, Robert D.  
; APPLICANT: Miller, Linda U.  
; TITLE OF INVENTION: Production of Host Cells Containing Multiple Integrating Vectors  
; FILE REFERENCE: GALA-08484  
; CURRENT APPLICATION NUMBER: US/10/759,315  
; PRIOR FILING DATE: 2004-01-16  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 13  
; LENGTH: 6255  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-759-315-13

Query Match 78.0%; Score 39.8; DB 18; Length 6255;  
Best Local Similarity 86.3%; Pred. No. 0.00025;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 5021 AGGATTGATCCTCGAATGGAATACTGTAATATGACCCGAAAGTTCCAGGCC 5071

RESULT 38  
US-10-947-881-13  
; Sequence 13, Application US/10947881  
; Publication No. US20050060762A1  
; GENERAL INFORMATION:  
; APPLICANT: Bleck, Gregory  
; TITLE OF INVENTION: Expression Vectors  
; FILE REFERENCE: GALA-06415  
; CURRENT APPLICATION NUMBER: US/10/947,881  
; CURRENT FILING DATE: 2004-09-23  
; PRIOR APPLICATION NUMBER: US/09/897,006  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 60/215,851  
; PRIOR FILING DATE: 2000-07-03  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13  
; LENGTH: 6255  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-947-881-13

Query Match 78.0%; Score 39.8; DB 19; Length 6255;  
Best Local Similarity 86.3%; Pred. No. 0.00025;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 5021 AGGATTGATCCTCGAATGGAATACTGTAATATGACCCGAAAGTTCCAGGCC 5071

RESULT 39  
US-11-018-895-13  
; Sequence 13, Application US/11018895  
; Publication No. US20050100952A1  
; GENERAL INFORMATION:  
; APPLICANT: Bremel, Robert  
; APPLICANT: Miller, Linda  
; APPLICANT: Bleck, Gregory  
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors  
; FILE REFERENCE: GALA-06416  
; CURRENT APPLICATION NUMBER: US/11/018,895  
; CURRENT FILING DATE: 2004-12-21  
; PRIOR APPLICATION NUMBER: US/09/897,511  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 60/215,925  
; PRIOR FILING DATE: 2000-07-03  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13  
; LENGTH: 6255  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-11-018-895-13

Query Match 78.0%; Score 39.8; DB 20; Length 6255;  
Best Local Similarity 86.3%; Pred. No. 0.00025;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCCGATATTCAGGCC 51  
Db 5021 AGGATTGATCCTCGAATGGAATACTGTAATATGACCCGAAAGTTCCAGGCC 5071

```

RESULT 40
US-10-239-656-66
; Sequence 66, Application US/10239656
; Publication No. US20040038339A1
; GENERAL INFORMATION:
; APPLICANT: KUPER, PETER
; APPLICANT: RIETHMULLER, GERT
; APPLICANT: LUTTERBUSE, RALF
; APPLICANT: BORSCHERT, KATRIN
; APPLICANT: KISCHEL, ROMAN
; APPLICANT: MAYER, MONIKA
; APPLICANT: HOFMEISTER, ROBERT
; TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE
; FILE REFERENCE: 029976/0106
; CURRENT APPLICATION NUMBER: US/10/239,656
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03414
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: EP 00106467.4
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic PS-10 single
; OTHER INFORMATION: chain Fv
US-10-239-656-66

Query Match      77.3%; Score 39.4; DB 17; Length 771;
Best Local Similarity 87.8%; Pred. No. 0.00022;
Matches 43; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      3 GATTGATCCTCGCAATCGTAAACTAAATATGACCCGATATTCAGGCC 51
Db      153 GATTGATCCTCGCAATCGTAAACTAAATATGACCCGATATTCAGGCC 201

Search completed: June 4, 2005, 18:27:21
Job time : 405.364 secs

```

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 12:51:52 ; Search time 2605.12 Seconds  
(without alignments)  
745.178 Million cell updates/sec

Title: US-10-089-452-25

Perfect score: 51

Sequence: 1 aagattgatctcggaatg.....atgacccgatattccaggcc 51

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_est3.\*

4: gb\_est4.\*

5: gb\_est5.\*

6: gb\_est6.\*

7: gb\_est7.\*

8: gb\_est8.\*

9: gb\_est9.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.4	81.2	320	2	BB870402
2	41.4	81.2	387	6	CA577790
3	41.4	81.2	524	4	BG965536
4	39.8	78.0	458	6	CA580083
5	39.8	78.0	496	6	CA578504
6	38.2	74.9	890	4	BG688002
7	37.8	74.1	709	4	BI1149624
8	37.8	74.1	884	2	BF164800
9	37.8	74.1	1057	2	BE306420
10	37.2	72.9	781	4	BG967007
11	36.2	71.0	928	2	BF143867
12	35.4	69.4	424	8	BH087049
13	31	60.8	877	2	BF159258
14	30.4	59.6	362	2	BB871319
15	27.2	53.3	531	4	BG959585
16	27.2	53.3	735	4	BG966898
17	27.2	53.3	862	2	BF128985
18	26.6	52.2	488	1	AA948436
19	26.6	52.2	646	6	CA681464
20	26.6	52.2	702	4	BG966479
21	26.6	52.2	895	4	BG685577
22	26.6	52.2	901	5	BQ709145
23	26.6	52.2	958	5	BQ523796
24	26.2	51.4	623	4	BG964955

25	26.2	51.4	798	4	BM007861
26	26.2	51.4	915	5	BQ883300
27	25.6	50.2	885	2	BF162852
28	25.6	50.2	916	5	BQ711164
29	25.4	49.8	633	4	BG340512
30	25.4	49.8	663	9	CC724490
31	25.4	49.8	699	9	CC615265
32	25.4	49.8	772	9	CC703985
33	25.4	49.8	774	9	CC615848
34	25.4	49.8	967	8	CC423318
35	25	49.0	209	7	T28938
36	25	49.0	320	5	BY346210
37	25	49.0	485	1	AJ548332
38	25	49.0	486	7	CR736049
39	25	49.0	488	1	AI791363
40	25	49.0	557	5	EX482986
41	25	49.0	645	5	EX099847
42	25	49.0	672	2	BF580337
43	25	49.0	698	2	BF584024
44	25	49.0	763	6	CB958770
45	25	49.0	815	4	BG756194

## ALIGNMENTS

RESULT 1  
BB870402

LOCUS

DEFINITION

lymph node Mus musculus

EST.

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB870402 320 bp mRNA linear EST 27-NOV-2001  
BB870402 RIKEN full-length enriched, adult male jejunal and colic  
lymph node Mus musculus cDNA clone G630021A01 5', mRNA sequence.  
BB870402 GI:17116612  
Mus musculus (house mouse)  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Akimura, T., Hirakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Inotani, K.,  
Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Konno, M.,  
Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,  
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
2001)  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001). Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## e mouse tissues.

## FEATURES

## source

```

1. 320
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G630021A01"
/sex="male"
/tissue_type="jejunal and colic lymph node"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult male jejunal
and colic lymph node"

```

## ORIGIN

```

Query Match      81.2%; Score 41.4; DB 2; Length 320;
Best Local Similarity 88.2%; Pred. No. 6.3e-05;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 1 AGAGTTGATCCTCGAATGCTAAATATGACCCGATATTCAGGCC 51
|
Db 254 AGGATTGATCCTCGAATGCTAAATATGACCCGATATTCAGGCC 304
|

```

## RESULT 2

## CA577790

## LOCUS

```

DEFINITION
K0714D10-5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)
CDNA Library (Long) Mus musculus cDNA clone NIA:K0714D10
IMAGE:30074925 5', mRNA sequence.

```

## CA577790

## EST.

```

CA577790.1 GI:25126181
Mus musculus (house mouse)

```

## ORGANISM

```

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

## REFERENCE

## AUTHORS

```

1 (bases 1 to 387)
Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Carter,M.G.,

```

## TITLE

```

Aiba,K., Taub,D., Longo,D.L., Keller,J. and Ko,M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)

```

## JOURNAL

## COMMENT

```

Unpublished (2001)
Other ESTs: K0714D10-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgun.gsc.nia.nih.gov
Plate: K0714 row: D column: 10
Seq primer: M13 Reverse
High quality sequence stop: 387
POLYA-No.

```

## FEATURES

## source

```

1. 387
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6NCR"
/db_xref="niaEST:K0714D10-5N"
/db_xref="taxon:10090"
/clone="NIA:K0714D10 IMAGE:30074925"
/tissue_type="Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-)"
/dev_stage="Age approx.10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)"
/notes="Vector: pSPORT1 (Invitrogen); Site_1: Sali; Site_2:

```

NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgun.gsc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were obtained from Drs. Dennis Taub, Dan Longo (National Institute on Aging, USA), Jonathan Keller (National Cancer Institute, USA). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen].

5'-pGACTAGTCTAGATCGGAGCGCGCTTTT-TTTT-3' from 0.9 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loxe-linker IL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with Sali and NotI enzymes and cloned into Sali/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.1 kb. The library was constructed by Yulan Piao (NIA)."

## ORIGIN

```

Query Match      81.2%; Score 41.4; DB 6; Length 387;
Best Local Similarity 88.2%; Pred. No. 6.5e-05;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

## Qy 1 AGAGTTGATCCTCGAATGCTAAATATGACCCGATATTCAGGCC 51

## Db 225 AGGATTGATCCTCGAATGCTAAATATGACCCGATATTCAGGCC 275

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

## |

Technologies. Note: this is a NCI\_CGAP Library."

```
ORIGIN
Query Match      81.2%; Score 41.4; DB 4; Length 524;
Best Local Similarity 88.2%; Pred. No. 6.7e-05;
Matches 45; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGAATGCTAAACCTAAATATGACCGCATATTCAGGGCC 51
    |||||
Db 238 AGGATTGATCTCGAATGCTAAACCTAAATATGACCGCATATTCAGGGCC 288

RESULT 4
CA580083
LOCUS
DEFINITION
CA580083.1 GI:25128474
CA580083
Mus musculus (house mouse)
EST.
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 458)
Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Carter,M.G.,
Aiba,K., Taub,D., Longo,D.L., Keller,J. and Ko,M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)
Unpublished (2001)
Other_ESTs: K0746F08-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0746 row: F column: 08
Seq primer: M13 Reverse
High quality sequence stop: 458
POLYA=No.

FEATURES
Location/Qualifiers
1..458
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6Ncr"
/db_xref="niaEST:K0746F08-5N"
/db_xref="taxon:10090"
/clone="NIA:K0746F08 IMAGE:30078019"
/tissue_type="Hematopoietic Stem Cell"
/dev_stage="Age approx.10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell"
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)"
/note="vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an Oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTTAGATCGGCGCGCCCTTTT-3'] from
0.9 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker L1-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
```

100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.1 kb. The library was constructed by Yulan Piao (NIA)."

```
ORIGIN
Query Match      78.0%; Score 39.8; DB 6; Length 458;
Best Local Similarity 86.3%; Pred. No. 0.00026;
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCTCGAATGCTAAACCTAAATATGACCGCATATTCAGGGCC 51
    |||||
Db 205 AGGATTGATCTCGAATGCTAAACCTAAATATGACCGCATATTCAGGGCC 255

RESULT 5
CA578504
LOCUS
DEFINITION
CA578504.1 GI:25126895
CA578504
Mus musculus (house mouse)
EST.
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 496)
Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Luo,A., Carter,M.G.,
Aiba,K., Taub,D., Longo,D.L., Keller,J. and Ko,M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)
Unpublished (2001)
Other_ESTs: K0724G04-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: K0724 row: G column: 04
Seq primer: M13 Reverse
High quality sequence stop: 496
POLYA=No.

FEATURES
Location/Qualifiers
1..496
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6Ncr"
/db_xref="niaEST:K0724G04-5N"
/db_xref="taxon:10090"
/clone="NIA:K0724G04 IMAGE:30075915"
/tissue_type="Hematopoietic Stem Cell"
(Lin-/c-Kit-/Sca-1-)
/dev_stage="Age approx.10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell"
(Lin-/c-Kit-/Sca-1-) cDNA Library (Long)"
/note="vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an Oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTTAGATCGGCGCGCCCTTTT-3'] from
0.9 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
```

to Lone-linker Ll-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.1 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 78.0%; Score 39.8; DB 6; Length 496;  
Best Local Similarity 86.3%; Pred. No. 0.00027;  
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCGATATTCAGGCC 51  
205 AGGATTGATCCTGAGGATGTTGAACTAAATATGCCCGGAATTCAGGCC 255

Db

RESULT 6  
LOCUS BG868002 890 bp mRNA linear EST 29-MAY-2001  
DEFINITION 602788352F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4919844 5', mRNA sequence.  
VERSION BG868002 1 GI:14218542  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 890)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLAM10835 row: g column: 13  
High quality sequence start: 2  
High quality sequence stop: 811.  
Location/Qualifiers  
1. .890  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4919844"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP\_SG2"  
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."

ORIGIN

Query Match 74.9%; Score 38.2; DB 4; Length 890;  
Best Local Similarity 84.3%; Pred. No. 0.0011;  
Matches 43; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGAATGTTAAACTAAATATGACCGATATTCAGGCC 51  
229 AGGATTGATCCTGAGGATGTTGAACTAAATATGCCCGGAATTCAGGCC 279

Db

RESULT 7

LOCUS BI149624

DEFINITION 602848436F1 NCI\_CGAP\_Lu29 Mus musculus cDNA clone IMAGE:5012001 5', mRNA sequence.

ACCESSION BI149624

VERSION BI149624.1 GI:14609625

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 709)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLAM11062 row: g column: 10  
High quality sequence stop: 705.  
Location/Qualifiers  
1. .709  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CZECH II"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5012001"  
/tissue\_type="spontaneous tumor, metastatic to mammary. Stem cell origin."  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Lu29"  
/notes="Organ: lung; Vector: pCMV-SPORT6; Site: 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 74.1%; Score 37.8; DB 4; Length 709;  
Best Local Similarity 85.7%; Pred. No. 0.0016;  
Matches 42; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 GATTGATCCTCGAATGTTAAACTAAATATGACCGATATTCAGGCC 51  
243 GATTGATCCTGAGGATGTTGAACTAAATATGCCCGGAATTCAGGCAC 291

Db

RESULT 8

LOCUS BF164800

DEFINITION 60177713F1 NCI\_CGAP\_Lu29 Mus musculus cDNA clone IMAGE:4019403 5', mRNA sequence.

ACCESSION BF164800

VERSION BF164800.1 GI:11045151

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 884)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.

ORIGIN

Query Match 74.1%; Score 37.8; DB 4; Length 709;  
Best Local Similarity 85.7%; Pred. No. 0.0016;  
Matches 42; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 GATTGATCCTCGAATGTTAAACTAAATATGACCGATATTCAGGCC 51  
243 GATTGATCCTGAGGATGTTGAACTAAATATGCCCGGAATTCAGGCAC 291

Db



CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1AM9271 row: i column: 04  
 High quality sequence stop: 727.  
 Location/Qualifiers

## FEATURES

source

1. .884

/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CZECH II"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4019403"  
 /tissue types="spontaneous tumor, metastatic to mammary.  
 Stem cell origin."  
 /lab host="DH10B"  
 /clone lib="NCI CGAP Lu29"  
 /note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: Sali;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"

## ORIGIN

Query Match 74.1%; Score 37.8; DB 2; Length 884;  
 Best Local Similarity 85.7%; Pred. No. 0.0016;  
 Matches 42; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 GATTGATCCTCGCAATGTTAAACTAAATATGACCCGATATTCAGGCC 51

|||||  
 Db 257 GATTGATCCTCGAGATGTTGAACTAAATATGCCCCGAGTTCAGGAC 305

RESULT 9  
 BE306420 1057 bp mRNA linear EST 26-OCT-2000  
 LOCUS 601103347F1 NCI\_CGAP\_Lu29 Mus musculus cDNA clone IMAGE:3495755 5',  
 DEFINITION mRNA sequence.

ACCESSION BE306420.1 GI:9159604

VERSION BE306420

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1057)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM8546 row: n column: 12

High quality sequence stop: 695.

Location/Qualifiers

1. .1057

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="IMAGE:3495755"

/tissue types="spontaneous tumor, metastatic to mammary.  
 Stem cell origin."

/lab host="DH10B"

/clone lib="NCI CGAP Lu29"

/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: Sali;

Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"

## ORIGIN

Query Match 74.1%; Score 37.8; DB 2; Length 1057;  
 Best Local Similarity 85.7%; Pred. No. 0.0016;  
 Matches 42; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 GATTGATCCTCGCAATGTTAAACTAAATATGACCCGATATTCAGGCC 51

|||||  
 Db 226 GATTGATCCTCGAGATGTTGAACTAAATATGCCCCGAGTTCAGGAC 274

## RESULT 10

BE967007

LOCUS 602834239F1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:498744 5',

DEFINITION mRNA sequence.

ACCESSION BE967007.1 GI:14354644

VERSION BE967007.1

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 781)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM11001 row: n column: 09

High quality sequence stop: 770.

Location/Qualifiers

1. .781

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:498744"

/lab host="DH10B (Ti phage-resistant)"

/clone lib="NCI\_CGAP\_Co24"

/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: Sali; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.6 kb. Constructed by Life

Technologies. Note: this is a NCI\_CGAP Library."

## ORIGIN

Query Match 72.9%; Score 37.2; DB 4; Length 781;  
 Best Local Similarity 84.0%; Pred. No. 0.0027;  
 Matches 42; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AAGATTGATCCTCGATGTTAAACTAAATATGACCCGATATTCAGGCC 50

|||||  
 Db 236 AGGATTGATCCTGAGACTGTCATATATATGACCCGAGTTCAGGCC 285

## RESULT 11

BF143867

LOCUS 601788273F1 NCI\_CGAP\_Lu30 Mus musculus cDNA clone IMAGE:4015795 5',

DEFINITION mRNA sequence.

ACCESSION BF143867

VERSION BF143867.1 GI:10982907

KEYWORDS EST.

```

SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9262 row: b column: 20
High quality sequence stop: 633.
FEATURES
source
1. .928
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4015795"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; transgenic model WNT-1, expression driven by
MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
ORIGIN
Query Match 71.0%; Score 36.2; DB 2; Length 928;
Best Local Similarity 83.7%; Pred. No. 0.0065;
Matches 41; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 3 GATTGATCTCGGAATGTTAACTAAATATGACCGATATTCAGGCC 51
|||||
Db 213 GATTGATCTCGGAATGTTGAGACTAAATATGTCCGGAAGTTCAGGAC 261
|||||

RESULT 12
BH087049 424 bp DNA linear GSS 18-JUL-2001
LOCUS RPCI-24-339A3.TV RPCI-24 Mus musculus genomic clone RPCI-24-339A3,
DEFINITION genomic survey sequence.
ACCESSION BH087049
VERSION BH087049.1 GI:14906954
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 424)
Zhao,S., Niernan,W., Malek,J., Shatsman,S., Akinret,B., Levine,M.,
Teagave,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-24
Other GSSs: RPCI-24-339A3.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong

SOURCE
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 928)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9262 row: b column: 20
High quality sequence stop: 633.
FEATURES
source
1. .424
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-339A3"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/notes="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI; The
RPCI-24 Mouse BAC library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."
ORIGIN
Query Match 69.4%; Score 35.4; DB 8; Length 424;
Best Local Similarity 86.7%; Pred. No. 0.012;
Matches 39; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 AGGATTGATCTCGGAATGTTAACTAAATATGACCGATATTC 45
|||||
Db 380 AGGATTGATCTCGGAATGTTGAACTAAATATGCCCCGAAATTC 424
|||||

RESULT 13
BH159258 877 bp mRNA linear EST 30-OCT-2000
LOCUS 601771226F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3990462 5',
DEFINITION mRNA sequence.
ACCESSION BH159258
VERSION BH159258.1 GI:11039356
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 877)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9202 row: c column: 07
High quality sequence stop: 684.
FEATURES
source
1. .877
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:3990462"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator

```

(pdjejong@mail.cho.org) . Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>) . BAC end page: [http://www.tigr.org/tadb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tadb/bac_ends/mouse/bac_end_intro.html) Plate: 339 row: A column: 3 Seq primer: T7 Class: BAC ends.

## FEATURES

source

```

1. .424
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-339A3"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/notes="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI; The
RPCI-24 Mouse BAC library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

```

## ORIGIN

```

Query Match 69.4%; Score 35.4; DB 8; Length 424;
Best Local Similarity 86.7%; Pred. No. 0.012;
Matches 39; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 AGGATTGATCTCGGAATGTTAACTAAATATGACCGATATTC 45
|||||
Db 380 AGGATTGATCTCGGAATGTTGAACTAAATATGCCCCGAAATTC 424
|||||

```

## RESULT 13

```

BH159258 877 bp mRNA linear EST 30-OCT-2000
LOCUS 601771226F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3990462 5',
DEFINITION mRNA sequence.
ACCESSION BH159258
VERSION BH159258.1 GI:11039356
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 877)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9202 row: c column: 07
High quality sequence stop: 684.

```

## REFERENCE

```

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9202 row: c column: 07
High quality sequence stop: 684.

```

## FEATURES

source

```

1. .877
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:3990462"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu29"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator

```



LOCUS  
DEFINITION 602834104F1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:4988593 5', linear EST 12-JUN-2000  
mRNA sequence.  
ACCESSION BG966898  
VERSION BG966898.1 GI:14354535  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
TITLE 1 (bases 1 to 735)  
JOURNAL NIH-MGC <http://mgc.nci.nih.gov/>.  
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Jeffrey F. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LHM11001 row: h column: 02  
High quality sequence stop: 728.  
Location/Qualifiers  
1. 735  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4988593"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Co24"  
/notes="Organ: colon; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
ORIGIN  
Query Match 53.3%; Score 27.2; DB 4; Length 735;  
Best Local Similarity 72.9%; Pred. No. 16;  
Matches 35; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
QY 2 AGATTGATCTCGCAATGGTAAACTAAATATGACCCGATATCCAGG 49  
|||||  
Db 235 AGATTATCTCGAAAGGTTTACTACTATACGCGCAAAATTCAGG 282  
|||||  
RESULT 17  
BF128985  
LOCUS  
DEFINITION BF128985 862 bp mRNA linear EST 24-OCT-2000  
mRNA sequence.  
ACCESSION BF128985  
VERSION BF128985.1 GI:10968025  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
TITLE 1 (bases 1 to 862)  
JOURNAL NIH-MGC <http://mgc.nci.nih.gov/>.  
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

```

from the same 3 libraries. The pools consisted of
1.M.A.C.E. clones 237480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match      52.2%; Score 26.6; DB 1; Length 488;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      3  GATTGATCCTCGCAATGTTAAATGTAATGACCCGATATTCAGGCC 51
      414  GATCAATGCTGGCAATGTACACACATATGCAAGATTCAGGC 366

Db

RESULT 19
CA681464/c
LOCUS      CA681464      646 bp      mRNA      linear      EST 24-NOV-2002
DEFINITION wlm24.pk0017.b2 wlm24 Triticum aestivum cDNA clone wlm24.pk0017.b2
5' end, mRNA sequence.
ACCESSION  CA681464
VERSION    CA681464.1  GI:25267145
KEYWORDS   EST.
SOURCE     Triticum aestivum (bread wheat)
ORGANISM   Triticum aestivum

REFERENCE
AUTHORS    Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
            Miao,G., Caraher,N. and Hanafey,M.K.
TITLE      DuPont Wheat cDNA Sequence
JOURNAL    Unpublished (2002)
COMMENT    Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.Tingey@USA.dupont.com
            Seq primer: M13.
            Location/Qualifiers
                1..646
                /organism="Triticum aestivum"
                /mol_type="mRNA"
                /cultivar="Stephens"
                /db_xref="taxon:4565"
                /clone="wlm24.pk0017.b2"
                /tissue_type="leaf"
                /clone_lib="wlm24"
                /notes="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
                XhoI; Wheat (Triticum aestivum L.) seedlings 24 hr after
                inoculation w/E. graminis"

FEATURES
source
1..646

ORIGIN
Query Match      52.2%; Score 26.6; DB 6; Length 646;
Best Local Similarity 70.0%; Pred. No. 26;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1  AAGATTGATCCTCGCAATGTTAAATGTAATGACCCGATATTCAGGC 50
      502  ACGATTAAATGTTGGGAATTTGNGAATTAATATCATCGATATCCCATGC 453

Db

RESULT 20
BG9666479
LOCUS      BG9666479      702 bp      mRNA      linear      EST 12-JUN-2001
DEFINITION 602832989F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4987319 5',
mRNA sequence.
ACCESSION  BG9666479
VERSION    BG9666479.1  GI:14354116
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)

```

/db\_xref="taxon:9606"  
/clone="IMAGE:4765212"  
/tissue type="primary B-cells from tonsils (cell line)"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NIH\_MGC\_48"  
/note="Organ: B-cells; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 52.2%; Score 26.6; DB 4; Length 895;  
Best Local Similarity 71.4%; Pred. No. 27;  
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 3 GATTGATCTCGCAATGTTAAACTAAATATGACCCGATATTCACGCC 51  
|||||  
Db 235 GATCAACGCTGCGCAATGTTAACTAAATATTCACCGAGTTCCAGGGC 283

## RESULT 22

BQ709145  
LOCUS BQ709145 901 bp mRNA linear EST 16-JUL-2002  
DEFINITION AGENCOURT 8475010 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:6301567  
5', mRNA sequence.

ACCESSION BQ709145

VERSION BQ709145.1

KEYWORDS GI:21848044

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 901)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov  
Tissue Procurement: Dr. Mark Watson  
cDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCW2517 row: k column: 08  
High quality sequence stop: 703.

## FEATURES

source

1..901  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6301567"  
/lab host="NIH\_MGC\_113"  
/clone lib="NIH\_MGC\_113"  
/note="Organ: spleen; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCACGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

## ORIGIN

Query Match 52.2%; Score 26.6; DB 5; Length 901;  
Best Local Similarity 71.4%; Pred. No. 27;  
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 3 GATTGATCTCGCAATGTTAAACTAAATATGACCCGATATTCACGCC 51  
|||||  
Db 241 GATCAGTCTCGCAATGTTAACTAAATATGCACAGACATTCCAGGGC 289

## RESULT 23

BUS23796  
LOCUS BUS23796 958 bp mRNA linear EST 13-SEP-2002  
DEFINITION AGENCOURT 10123257 NCI\_CGAP\_Co24 Mus musculus cDNA clone  
IMAGE:6530504 5', mRNA sequence.

ACCESSION BUS23796

VERSION BUS23796.1

KEYWORDS GI:22834235

SOURCE EST.

ORGANISM Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 958)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

## REFERENCE

AUTHORS NIH-MGC

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM14132 row: f column: 08  
High quality sequence stop: 608.

## FEATURES

source

1..958  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6530504"  
/lab host="NIH\_MGC\_Co24"  
/clone lib="NCI\_CGAP\_Co24"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: Sall; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

## ORIGIN

Query Match 52.2%; Score 26.6; DB 5; Length 958;  
Best Local Similarity 71.4%; Pred. No. 27;  
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 3 GATTGATCTCGCAATGTTAAACTAAATATGACCCGATATTCACGCC 51  
|||||  
Db 245 GATTATCTCGCAATGTTAAATATTAATACATGAGAGTTCCAGGGC 293

## RESULT 24

BG964955  
LOCUS BG964955 623 bp mRNA linear EST 12-JUN-2001  
DEFINITION 602829272F1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:4984169 5',  
mRNA sequence.

ACCESSION BG964955

VERSION BG964955.1

KEYWORDS GI:14352592

SOURCE EST.

ORGANISM Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 623)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LHAM10989 row: o column: 18  
 High quality sequence stop: 617.  
 Location/Qualifiers  
 1..623  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4984169"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Co24"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI\_CGAP Library."

## ORIGIN

Query Match 51.4%; Score 26.2; DB 4; Length 623;  
 Best Local Similarity 72.3%; Pred. No. 37;  
 Matches 34; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 5 TTGATCTCGCAATGGTAAACTAAATATGACCGCATATTCAGGCC 51  
 |||||  
 Db 239 TTAATCTCGAATGGTGATACCTACACAGAAATTCGAGGCC 285

## RESULT 25

BM007861 798 bp mRNA linear EST 30-OCT-2001  
 LOCUS 603617556f1 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:5450749 5',  
 DEFINITION mRNA sequence.  
 BM007861  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Dr. Mark Watson  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LHAM1939 row: p column: 14  
 High quality sequence stop: 798.  
 Location/Qualifiers  
 1..798  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5450749"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_113"  
 /note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor:

FEATURES  
source

Query Match 51.4%; Score 26.2; DB 5; Length 915;  
 Best Local Similarity 72.3%; Pred. No. 38;  
 Matches 34; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 5 TTGATCTCGCAATGGTAAACTAAATATGACCGCATATTCAGGCC 51  
 |||||  
 Db 256 TTGATCTCGAATGGTGAACAAGTTATCGACGAGAAATTCAGGCC 302

GGCACCAG (G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 51.4%; Score 26.2; DB 4; Length 798;  
 Best Local Similarity 72.3%; Pred. No. 38;  
 Matches 34; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 5 TTGATCTCGCAATGGTAAACTAAATATGACCGCATATTCAGGCC 51  
 |||||  
 Db 252 TTGATCTCGAATGGTGAACAAGTTATCGACGAGAAATTCAGGCC 298

## RESULT 26

BM0083300 915 bp mRNA linear EST 16-AUG-2002  
 LOCUS 8122430 Lupski\_dorsal\_root\_ganglion Homo sapiens cDNA  
 DEFINITION AGENCOURT 8122430 Lupski\_dorsal\_root\_ganglion Homo sapiens cDNA  
 clone IMAGE:6181650 5', mRNA sequence.  
 BM0083300  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LHAM13566 row: f column: 19  
 High quality sequence stop: 564.  
 Location/Qualifiers  
 1..915  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6181650"  
 /sex="male"  
 /tissue\_type="dorsal root ganglia"  
 /dev\_stage="adult, 36 yr"  
 /lab\_host="DH10B"  
 /clone\_lib="Lupski dorsal root ganglion"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1: NotI; Site 2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors:  
 5'-TCGACCCAGCGTCCG-3' and  
 5'-GACTAGTCTAGATCGGAGCGGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

FEATURES  
source

Query Match 51.4%; Score 26.2; DB 5; Length 915;  
 Best Local Similarity 72.3%; Pred. No. 38;  
 Matches 34; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 5 TTGATCTCGCAATGGTAAACTAAATATGACCGCATATTCAGGCC 51  
 |||||  
 Db 256 TTGATCTCGAATGGTGAACAAGTTATCGACGAGAAATTCAGGCC 302

```
RESULT 27
BF162852          885 bp  mRNA  linear  EST 30-OCT-2000
LOCUS             601767129F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3982950 5',
DEFINITION        mRNA sequence.
ACCESSION         BF162852
VERSION           BF162852.1 GI:11043083
KEYWORDS          EST.
SOURCE            Mus musculus (house mouse)
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE         1 (bases 1 to 885)
AUTHORS           NIH-MGC http://mgc.nci.nih.gov/.
TITLE            National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgabbs-r@mail.nih.gov
                  Tissue Procurement: Gilbert Smith, Ph.D.
                  CDNA Library Preparation: Life Technologies, Inc.
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: L1AM9182 row: j column: 07
                  High quality sequence stop: 667.
FEATURES          source
                  1..885
                     Location/Qualifiers
                        /organism="Mus musculus"
                        /mol_type="mRNA"
                        /strain="CZECH II"
                        /db_xref="taxon:10090"
                        /clone="IMAGE:3982950"
                        /tissue_type="spontaneous tumor, metastatic to mammary.
                        Stem cell origin."
                        /lab_host="DH10B"
                        /clone_lib="NCI CGAP Lu29"
                        /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: Salt;
                        Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                        Library constructed by Life Technologies. Investigator
                        providing samples: Gilbert Smith, NIH"
ORIGIN
Query Match      50.2%; Score 25.6; DB 2; Length 885;
Best Local Similarity 70.8%; Pred. No. 65;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 4 ATTGATCCTCGGAATGCTAAACTAAATATGACCCGATATTCAGGCC 51
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 217 ATTATCTCGGAATAGTGATCTACTACTACACAGAGTTCAAGGCC 264
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 28
BQ711164
LOCUS             AGENCOURT_7977133 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6215982
DEFINITION        5', mRNA sequence.
ACCESSION         BQ711164
VERSION           BQ711164.1 GI:21850063
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE         1 (bases 1 to 916)
AUTHORS           NIH-MGC http://mgc.nci.nih.gov/.
TITLE            National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgabbs-r@mail.nih.gov
```

```
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2385 row: m column: 07
High quality sequence stop: 428.
FEATURES          source
                  1..916
                     Location/Qualifiers
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:6215982"
                        /lab_host="DH10B (phage-resistant)"
                        /clone_lib="NIH MGC 113"
                        /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
                        EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                        into EcoRI/XhoI sites using the following 5' adaptor:
                        GGCACGAG(G). Library constructed by Ling Hong in the
                        laboratory of Gerald M. Rubin (University of California,
                        Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                        Superscript II RT (Life Technologies). Note: this is a
                        NIH_MGC Library."
ORIGIN
Query Match      50.2%; Score 25.6; DB 5; Length 916;
Best Local Similarity 70.8%; Pred. No. 65;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 4 ATTGATCCTCGGAATGCTAAACTAAATATGACCCGATATTCAGGCC 51
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 150 ATTGATCTCGAGATGCTGAACAATATACGCACAGAGTTGCAGGCC 197
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 29
BG340512
LOCUS             602462069F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4575064 5',
DEFINITION        mRNA sequence.
ACCESSION         BG340512
VERSION           BG340512.1 GI:13146950
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE         1 (bases 1 to 633)
AUTHORS           NIH-MGC http://mgc.nci.nih.gov/.
TITLE            National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgabbs-r@mail.nih.gov
                  Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
                  CDNA Library Preparation: Ling Hong/Rubin Laboratory
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: L1CM1285 row: i column: 17
                  High quality sequence stop: 630.
FEATURES          source
                  1..633
                     Location/Qualifiers
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:4575064"
                        /tissue_type="primary B-cells from tonsils (cell line)"
                        /lab_host="DH10B (phage-resistant)"
                        /clone_lib="NIH MGC 48"
                        /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
```



Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGACAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by Ling  
 Hong in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 49.8%; Score 25.4; DB 4; Length 633;  
 Best Local Similarity 68.6%; Pred. No. 74;  
 Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 AAGTTGATCTCGAATGTAATACTAATATGACCGGATATTCAGGCG 51  
 |||||  
 Db 217 AGGTTTGATCTCGAATGTAATACTAATATGACCGGATATTCAGGCG 267

## RESULT 30

CC724490/c  
 LOCUS  
 DEFINITION OGVEM75TH ZM 0.7\_1.5\_KB Zea mays genomic clone ZMMBma0537M05,  
 genomic survey sequence.

ACCESSION CC724490.1 GI:32143423

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## Ze mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

1 (bases 1 to 663)  
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.  
 Consortium for Maize Genomics  
 Unpublished (2002)  
 Other\_GSSs: OGVEM75TV  
 Contact: Cathy Whitelaw  
 TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TR  
 Class: sheared ends.

1. .663  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBma0537M05"  
 /clone\_lib="ZM 0.7\_1.5\_KB"  
 /note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

## FEATURES

## source

Query Match 49.8%; Score 25.4; DB 9; Length 663;  
 Best Local Similarity 74.4%; Pred. No. 75;  
 Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

## ORIGIN

QY 8 ATCTCGAATGTAATACTAATATGACCGGATATTCAGGCG 50  
 |||||  
 Db 209 ATCATATTAACTGTAATACTAATATGACCTCGCTCTTGCAGCG 167

## RESULT 31

CC615265/c  
 LOCUS  
 DEFINITION OGVEM063TH ZM 0.7\_1.5\_KB Zea mays genomic clone ZMMBma0537L05,  
 genomic survey sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## Ze mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

1 (bases 1 to 699)  
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.  
 Consortium for Maize Genomics  
 Unpublished (2002)  
 Other\_GSSs: OGVEM063TV  
 Contact: Cathy Whitelaw  
 TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TR  
 Class: sheared ends.

1. .699  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBma0537L05"  
 /clone\_lib="ZM 0.7\_1.5\_KB"  
 /note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

## FEATURES

## source

Query Match 49.8%; Score 25.4; DB 9; Length 699;  
 Best Local Similarity 74.4%; Pred. No. 75;  
 Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

## ORIGIN

## QY

## Db

## RESULT 32

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## Ze mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

8 ATCTCGAATGTAATACTAATATGACCGGATATTCAGGCG 50  
 |||||  
 Db 209 ATCATATTAACTGTAATACTAATATGACCTCGCTCTTGCAGCG 167

9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TR  
 Class: sheared ends.

1. .663  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBma0537L05"  
 /clone\_lib="ZM 0.7\_1.5\_KB"  
 /note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

Query Match 49.8%; Score 25.4; DB 9; Length 663;  
 Best Local Similarity 74.4%; Pred. No. 75;  
 Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 8 ATCTCGAATGTAATACTAATATGACCGGATATTCAGGCG 50  
 |||||  
 Db 209 ATCATATTAACTGTAATACTAATATGACCTCGCTCTTGCAGCG 167

9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TR  
 Class: sheared ends.

1. .699  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBma0537L05"  
 /clone\_lib="ZM 0.7\_1.5\_KB"  
 /note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

Query Match 49.8%; Score 25.4; DB 9; Length 663;  
 Best Local Similarity 74.4%; Pred. No. 75;  
 Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 8 ATCTCGAATGTAATACTAATATGACCGGATATTCAGGCG 50  
 |||||  
 Db 209 ATCATATTAACTGTAATACTAATATGACCTCGCTCTTGCAGCG 167

9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TR  
 Class: sheared ends.

1. .699  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBma0537L05"  
 /clone\_lib="ZM 0.7\_1.5\_KB"  
 /note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

```

source
1. .772
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0537L06"
/clone_lib="ZM 0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 49.8%; Score 25.4; DB 9; Length 772;
Best Local Similarity 74.4%; Pred. No. 76;
Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 8 ATCTCGGATGTGTAACCTAAATATGACCCGATATTCAGGC 50
|||||
Db 209 ATCATATTACCGTAAACTAAATATGACTCGCTCTTGCACGC 167

RESULT 33
CC615848/c
LOCUS
DEFINITION
CGV064TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0537L07,
genomic survey sequence.
CC615848
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays

REFERENCE
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 774)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nummer,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGV064TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1. .774
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0537L07"
/clone_lib="ZM 0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 49.8%; Score 25.4; DB 9; Length 774;
Best Local Similarity 74.4%; Pred. No. 76;
Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 8 ATCTCGGATGTGTAACCTAAATATGACCCGATATTCAGGC 50
|||||
Db 209 ATCATATTACCGTAAACTAAATATGACTCGCTCTTGCACGC 167

RESULT 34
CC423318/c
LOCUS
DEFINITION
CGV064TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0537L07,
genomic survey sequence.
CC423318
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays

REFERENCE
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 967)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUHBL02TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1. .967
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma413A04"
/clone_lib="ZM 0.6_1.0_KB"
/notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

COT selected genomic DNA library"

ORIGIN
Query Match 49.8%; Score 25.4; DB 8; Length 967;
Best Local Similarity 74.4%; Pred. No. 78;
Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 8 ATCTCGGATGTGTAACCTAAATATGACCCGATATTCAGGC 50
|||||
Db 814 ATCATATTACCGTAAACTAAATATGACTCGCTCTTGCACGC 772

RESULT 35
T28938/c
LOCUS
DEFINITION
T28938 Human White blood cells Homo sapiens cDNA 5' end similar
to immunoglobulin mu heavy chain, VDJ regions (GB:M17751)
(HT:3055), mRNA sequence.
T28938
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 209)
Adam,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,
Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitch,H.W.M., Fritchman,J.L., Geoghagen,N.S.M.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,J.P.S.,
Kelley,J.M., Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M.,
Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T.,
Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,
Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A.,
Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferrie,A.,
Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,

```

Meisner, P.S., Olsen, H., Raymond, L., Wei, Y.-F., Wing, J., Xu, C., Yu, G.-L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseitine, W.A., Fleide, C., Fraser, C.M. and Venter, J.C.  
Initial Assessment of Human Gene Diversity and Expression Patterns Based Upon 83 Million Basepairs of cDNA Sequence  
Nature 377, 3-174 (1995)  
96026280  
7566098

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Contact: Venter, JC  
The Institute for Genomic Research  
932 Clopper Rd, Gaithersburg, MD 20878  
Tel: 3018699056  
Fax: 3018699423  
Email: tdbinfo@tdb.tigr.org

For clone availability, additional sequence and expression information related to this EST, please contact the TIGR Database (tdbinfo@tdb.tigr.org)

Seq primer: M13 Reverse.

## FEATURES

## source

Location/Qualifiers  
1..209  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="ATCC (inhost):104674"  
/db\_xref="taxon:9606"  
/tissue types="white blood cells"  
/clone\_lib="Human white blood cells"  
/note="Organ: blood"

## ORIGIN

Query Match 49.0%; Score 25; DB 7; Length 209;  
Best Local Similarity 69.4%; Pred. No. 95;  
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 3 GATTGATCTCGCAATGTTAAACATAATATGACCCGATATTCAGGCC 51

Db 77 GATCAAGCTGGCAATGTTAAACATAATATTTACAGAGTTCAGGCC 29

## RESULT 36

## BY346210

## LOCUS

DEFINITION BY346210 RIKEN full-length enriched, whole joints Mus musculus cDNA clone L730022G23 5', mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 320)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisler, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierzki, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyszewski-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Miyazaki, A., Sakai, K., Saeki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
22354683

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Vassilis Aidinis (Biomedical Sciences Research Center 'Al. Fleming' Institute of Immunology 14-16 Al. Fleming street 16672 Vari, Greece) whose assistance we gratefully acknowledge.

Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

## FEATURES

## source

Location/Qualifiers  
1..320  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="L730022G23"  
/tissue types="whole joints"  
/clone\_lib="RIKEN full-length enriched, whole joints"

## ORIGIN

Query Match 49.0%; Score 25; DB 5; Length 320;  
Best Local Similarity 69.4%; Pred. No. 99;  
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 3 GATTGATCTCGCAATGTTAAACATAATATGACCCGATATTCAGGCC 51

Db 254 GATTATCTCTGCAAGCGGTAACTAAGTACAAATGAGNAGTTCAGGCC 302

## RESULT 37

## AJ548332

## LOCUS

DEFINITION AJ548332 MTAPHFU Medicago truncatula cDNA clone mtaehac110007d09, mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

AJ548332.1 GI:28612046 EST.





*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 10:28:02 ; Search time 747.667 Seconds  
(without alignments)  
2138.683 Million cell updates/sec

Title: US-10-089-452-26

Perfect score: 33

Sequence: 1 ccattattacgtagtctcgttggtttac 33

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hhg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	33	6	AX113451 Sequence
2	33	100.0	33	6	AX113586 Sequence
3	33	100.0	33	6	AX816253 Sequence
4	33	100.0	360	6	AX113428 Sequence
5	33	100.0	360	6	AX113563 Sequence
6	33	100.0	360	6	AX816230 Sequence
7	24	72.7	296	10	AF458198 Mus muscu
8	23.4	70.9	383	10	AY436976 Mus muscu
9	23.2	70.3	219	10	AY182679 Mus muscu
10	23.2	70.3	357	10	MMV32A Z12814 M.musculus
11	23	69.7	345	6	E16378 cDNA encodi
12	23	69.7	246199	2	AC125843 Rattus no
13	23	69.7	251052	2	AC130129 Rattus no
14	22.8	69.1	375	10	AB058179 Mus muscu
15	22.4	67.9	333	10	AF072776 Mus muscu
16	22	66.7	121360	2	AC143115 Macaca mu
17	22	66.7	183472	2	AC132625 Mus muscu
18	22	66.7	262190	10	AC131673 Mus muscu
19	21.8	66.1	388	10	AY454493 Mus muscu

C	20	21.8	66.1	246951	2	AC107248	AC107248 Rattus no
	21	21.6	65.5	37	10	MUSIGDVAJ	L36732 Mus musculu
	22	21.6	65.5	162	10	AY182575	AY182575 Mus muscu
	23	21.6	65.5	330	10	AY182599	AY182599 Mus muscu
	24	21.6	65.5	198284	2	CR478208	CR478208 Danio rer
	25	21.4	64.8	161671	2	AC018369	AC018369 Homo sapi
	26	21.4	64.8	161731	9	AC107021	AC107021 Homo sapi
	27	21.4	64.8	225551	2	AC150560	AC150560 Mus muscu
	28	21.2	64.2	229231	2	AC096139	AC096139 Rattus no
	29	21.2	64.2	250348	2	AC108553	AC108553 Rattus no
	30	21.2	64.2	272952	2	AC109685	AC109685 Rattus no
	31	21	63.6	50	10	AF071679	AF071679 Mus muscu
	32	21	63.6	13078	1	AE008602	AE008602 Rickettsi
	33	21	63.6	173373	9	AL356968	AL356968 Human DNA
	34	21	63.6	184279	10	AC140393	AC140393 Mus muscu
	35	21	63.6	216153	10	AL627183	AL627183 Mouse DNA
	36	21	63.6	223642	2	AC108290	AC108290 Rattus no
	37	21	63.6	236824	2	AC122632	AC122632 Rattus no
	38	21	63.6	240178	2	AC094173	AC094173 Rattus no
	39	21	63.6	246933	2	AC096662	AC096662 Rattus no
	40	21	63.6	246975	2	AC130059	AC130059 Rattus no
	41	21	63.6	255619	2	AC107503	AC107503 Rattus no
	42	21	63.6	302893	2	AC111936	AC111936 Rattus no
	43	20.8	63.0	318	10	AY369864	AY369864 Mus muscu
	44	20.8	63.0	384	10	AY436951	AY436951 Mus muscu
	45	20.8	63.0	1611	1	AF237938	AF237938 Pasteurel

#### ALIGNMENTS

RESULT 1

AX113451

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

mbH (DE)

Location/Qualifiers

1. .33

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="CDR"

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

SOURCE  
ORGANISM  
synthetic construct  
other sequences; artificial sequences.

REFERENCE  
1  
AUTHORS  
Reiter,C., Cullmann,G., Heppner,P., Ringeis,A., Mueller,H. and Haindl,E.  
TITLE  
Improved method for the detection of acid resistant microorganisms in a stool  
JOURNAL  
Patent: WO 0127613-A 26 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung (DE)

FEATURES  
source  
Location/Qualifiers  
1. .33  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="CDR"

ORIGIN  
Query Match 100.0%; Score 33; DB 6; Length 33;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCATTATTACGCTAGTTCCTGCTTGGCTTAC 33  
|||||  
Db 1 CCCATTATTACGCTAGTTCCTGCTTGGCTTAC 33

RESULT 3  
AX816253  
LOCUS  
AX816253  
DEFINITION  
Sequence 26 from Patent EP1336850.  
ACCESSION  
AX816253  
VERSION  
AX816253.1 GI:39646775  
KEYWORDS  
synthetic construct  
ORGANISM  
synthetic construct  
other sequences; artificial sequences.

REFERENCE  
1  
AUTHORS  
Reiter,C., Cullmann,G., Mueller,H., Heppner,P., Haindl,E. and Ringeis,A.  
TITLE  
Improved method for the detection of acid resistant microorganisms in a stool  
JOURNAL  
Patent: EP 1336850-A 26 20-AUG-2003;  
Connex Gesellschaft zur Optimierung von Forschung und Ent wicklung (DE)

FEATURES  
source  
Location/Qualifiers  
1. .33  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 100.0%; Score 33; DB 6; Length 33;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCATTATTACGCTAGTTCCTGCTTGGCTTAC 33  
|||||  
Db 1 CCCATTATTACGCTAGTTCCTGCTTGGCTTAC 33

RESULT 4  
AX113428  
LOCUS  
AX113428  
DEFINITION  
Sequence 3 from Patent WO0127612.  
ACCESSION  
AX113428  
VERSION  
AX113428.1 GI:13939696  
KEYWORDS  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
1  
AUTHORS  
Reiter,C., Cullmann,G., Lakner,M., Truee,A., Dehnert,S. and Schwartz,G.  
TITLE  
Immuno-chromatographic rapid assay in order to detect acid-resistant microorganisms in the stool  
JOURNAL  
Patent: WO 0127612-A 3 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung mbH (DE)

FEATURES  
source  
Location/Qualifiers  
1. .360  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"

ORIGIN  
Query Match 100.0%; Score 33; DB 6; Length 360;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCATTATTACGCTAGTTCCTGCTTGGCTTAC 33  
|||||  
Db 295 CCCATTATTACGCTAGTTCCTGCTTGGCTTAC 327

RESULT 5  
AX113563  
LOCUS  
AX113563  
DEFINITION  
Sequence 3 from Patent WO0127613.  
ACCESSION  
AX113563  
VERSION  
AX113563.1 GI:13939767  
KEYWORDS  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
1  
AUTHORS  
Reiter,C., Cullmann,G., Heppner,P., Ringeis,A., Mueller,H. and Haindl,E.  
TITLE  
Improved method for the detection of acid resistant microorganisms in a stool  
JOURNAL  
Patent: WO 0127613-A 3 19-APR-2001;  
Connex Gesellschaft zur Optimierung von Forschung und Entwicklung (DE)

FEATURES  
source  
Location/Qualifiers  
1. .360  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"

ORIGIN  
Query Match 100.0%; Score 33; DB 6; Length 360;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCATTATTACGCTAGTTCCTGCTTGGCTTAC 33  
|||||  
Db 295 CCCATTATTACGCTAGTTCCTGCTTGGCTTAC 327

RESULT 6  
AX816230  
LOCUS  
AX816230  
DEFINITION  
Sequence 3 from Patent EP1336850.  
ACCESSION  
AX816230  
VERSION  
AX816230.1 GI:39646761  
KEYWORDS  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
1  
AUTHORS  
Reiter,C., Cullmann,G., Mueller,H., Heppner,P., Haindl,E. and Ringeis,A.



**TITLE** Improved method for the detection of acid resistant microorganisms in a stool  
**JOURNAL** Patent: EP 1336950-A 3 20-AUG-2003;  
 Connex Gesellschaft zur Optimierung von Forschung und Ent wicklung (DE)

**FEATURES**  
 source Location/Qualifiers  
 1..360  
 /organism="Mus musculus"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:10090"

**ORIGIN**

Query Match 100.0%; Score 33; DB 6; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCATTATTACGCTAGTCTCTGTTGCTTAC 33  
 |||||  
 Db 295 CCCATTATTACGCTAGTCTCTGTTGCTTAC 327

**RESULT 7**  
 AF458198 296 bp DNA linear ROD 07-MAR-2003  
**LOCUS** Mus musculus clone B6F3MH160(3) immunoglobulin heavy chain VDJ gene, partial cds.  
**ACCESSION** AF458198  
**VERSION** AF458198.1 GI:28875211  
**KEYWORDS**  
**SOURCE** Mus musculus (house mouse)  
**ORGANISM**

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Chang, S., Hsu, K. and Mohan, C.  
**REFERENCE** Aberrant ligand-based selection of naive peripheral B-cells in the context of an autoimmune genome  
**TITLE** Unpublished  
**JOURNAL** 2 (bases 1 to 296)  
**REFERENCE** Chang, S., Hsu, K. and Mohan, C.  
**AUTHORS** Direct Submission  
**TITLE** Submitted (05-DEC-2001) Internal Medicine Rheumatology, UT  
**JOURNAL** Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390, USA

**FEATURES**  
 source Location/Qualifiers  
 1..296  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /chromosome="12"  
 /clone="B6F3MH160(3)"  
 /cell\_type="follicular B-cell"  
 /tissue\_type="spleen"  
 /rearranged  
 <1..>296  
 /product="immunoglobulin heavy chain VDJ"  
 <1..>296

**mRNA**  
**CDS**  
 /product="immunoglobulin heavy chain VDJ"  
 /product="immunoglobulin heavy chain VDJ"  
 /protein\_id="AA059820.1"  
 /db\_xref="GI:28875212"  
 /translation="SVTLSCASGYFTDYEMHWKQTPVHGLEWIGAIIDPETGGTAY NQRFKGAAILTADKSSSTAYMELSLTSDSAAYICTNYGSTWTFAYWGQTLV"

**ORIGIN**

Query Match 72.7%; Score 24; DB 10; Length 296;  
 Best Local Similarity 84.4%; Pred. No. 11;  
 Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 CCATTATTACGCTAGTCTCTGTTGCTTAC 33  
 |||||  
 Db 243 CAAATTACTAGGCTAGTCTCTGTTGCTTAC 274

**RESULT 8**

AY436976 383 bp DNA linear ROD 15-MAR-2004  
**LOCUS** Mus musculus ZDC1 ANA immunoglobulin heavy chain gene, partial cds.  
**DEFINITION** AY436976  
**ACCESSION** AY436976  
**VERSION** AY436976.1 GI:40794900  
**KEYWORDS**  
**SOURCE** Mus musculus (house mouse)  
**ORGANISM**

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Liang, Z., Xie, C., Chen, C., Kreska, D., Hsu, K., Li, L., Zhou, X. J. and Mohan, C.  
**REFERENCE** Pathogenic Profiles and Molecular Signatures of Antinuclear Autoantibodies Rescued from NZM2410 Lupus Mice  
**AUTHORS** J. Exp. Med. 199 (3), 381-398 (2004)  
 14757744  
 2 (bases 1 to 383)  
**REFERENCE** Liang, Z., Xie, C., Chen, C., Kreska, D., Hsu, K., Li, L., Zhou, J. X. and Mohan, C.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (29-SEP-2003) Simmons Arthritis Research Center and Center for Immunology, University of Texas Southwestern Medical School, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA

**FEATURES**  
 source Location/Qualifiers  
 1..383  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="NZM2410"  
 /db\_xref="taxon:10090"  
 /chromosome="12"  
 /cell\_line="ZDC1"  
 /cell\_type="spleen plasma B-cell hybridoma"  
 /rearranged  
 <1..>383  
 /product="ANA immunoglobulin heavy chain"  
 <1..>383  
 /notes="lupus; anti-DNA; glomerulonephritis"  
 /codon\_start=3  
 /product="ANA immunoglobulin heavy chain"  
 /protein\_id="AA891004.1"  
 /db\_xref="GI:40794901"  
 /translation="ELVRFGASVYKMSCTAGFNINDYMHVQRPQGLEWIGWIDP ANGDEYASKFQKATITATSSNTAYLQLSSLTSDTAVYCSPHYGSTWTFAYWGQ GTLVTVSASQSPNVFPLVSCSP"

**mRNA**  
**CDS**  
 /product="ANA immunoglobulin heavy chain"  
 /notes="lupus; anti-DNA; glomerulonephritis"  
 /codon\_start=3  
 /product="ANA immunoglobulin heavy chain"  
 /protein\_id="AA891004.1"  
 /db\_xref="GI:40794901"  
 /translation="ELVRFGASVYKMSCTAGFNINDYMHVQRPQGLEWIGWIDP ANGDEYASKFQKATITATSSNTAYLQLSSLTSDTAVYCSPHYGSTWTFAYWGQ GTLVTVSASQSPNVFPLVSCSP"

**ORIGIN**

Query Match 70.9%; Score 23.4; DB 10; Length 383;  
 Best Local Similarity 81.8%; Pred. No. 20;  
 Matches 27; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1 CCCATTATTACGCTAGTCTCTGTTGCTTAC 33  
 |||||  
 Db 267 CCTCAATTACTACGCTAGTACCTGCTTGCTTAC 299

**RESULT 9**  
 AY182679 219 bp mRNA linear ROD 13-JAN-2003  
**LOCUS** Mus musculus clone L2FLiDC-P18 immunoglobulin heavy chain variable region mRNA, partial cds.  
**DEFINITION** AY182679  
**ACCESSION** AY182679.1 GI:27729487  
**VERSION** AY182679.1 GI:27729487  
**KEYWORDS** Mus musculus (house mouse)  
**SOURCE** Mus musculus  
**ORGANISM**

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Kretschmer, K., Engel, H. and Weiss, S.  
**REFERENCE** 1 (bases 1 to 219)

**AUTHORS**

TITLE	Strong antigenic selection shaping the immunoglobulin heavy chain repertoire of B-1a lymphocytes in lambda 2(315) transgenic mice
JOURNAL	Eur. J. Immunol. 32 (8), 2317-2327 (2002)
MEDLINE	22197943
PUBMED	12209645
PREFERENCE	2 (bases 1 to 219)
AUTHORS	Kretschmer, K., Engel, H. and Weiss, S.
TITLE	Direct Submission
JOURNAL	Submitted (18-NOV-2002) Molecular Immunology, GBF, German Research Centre for Biotechnology, Mascheroder Weg 1, Braunschweig 38124,

FEATURES	Location/Qualifiers
source	1. .219

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/isolation_source="in vitro differentiated liver"
/db_xref="taxon:10090"
/clone="L2FLidC-P18"
/cell_type="B lymphocyte"
/dev_stage="fetus"
/note="lambda2 MOPC315 transgenic L2 mouse"
<1..>219
/codon_start=1
/product="immunoglobulin heavy chain variable region"
/protein_id="AA019035.1"
/db_xref="GI:27729488"
/translation="LVAINSDGGSTYYPDMMERRFIISDNTKTKLYLQWSSLRSD
TALYVCARHYGSGTWFAWYGQGLVTVSA"

```

## ORIGIN

Query Match	70.3%	Score 23.2;	DB 10;	Length 219;
Best Local Similarity	89.3%	Pred. No. 25;		
Matches	25;	Conservative	0; Mismatches	3; Indels 0; Gaps 0;
QY	6	TTATTACGCTAGTTCCTGGTTTGCTTAC	33	
Db	159	TTACTACGGTAGTACCTGGTTTGCTTAC	186	

## RESULT 10

MMV132A	MMV32A	357 bp	mRNA	linear	ROD 23-JUN-1992
LOCUS	M.musculus mRNA for VH-gen sequence of naturally occurring, somatically mutated memory B cell.				
DEFINITION	212814 X59730				
ACCESSION	212814.1 GI:55199				
VERSION	somatic mutation.				
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 357)				
AUTHORS	Schitteck, B. and Rajewsky, K.				
TITLE	Natural occurrence and origin of somatically mutated memory B cells				

JOURNAL  
REFERENCE  
2 (bases 1 to 35)  
in mice  
J. Exp. Med. (1992)

**AUTHORS** Schitteck, B.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (07-JUN-1992)  
**EDITORIAL** B. Schitteck, Institut of Genetics, Wevertal

**FEATURES**  
**source**

```

/organism="Mus musculus"
/mol_type="rRNA"
/strain="C57BL/6"
/isolate="V3_2A"
/db_xref="taxon:10090"
/clone="V3_2A"
/haplotype="IgHb"
/sex="female"
/cell_type="B-cell"

```

## ORIGIN

Query Match	70.3%	Score 23.2;	DB 10;	Length 357;
Best Local Similarity	89.3%;	Pred. No. 25;		
Matches 25;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	6	TTATTACGTAGTTCCTGGTTTGCTTAC	33	
Db	297	TTACTACGTAGTTCCTGGTTTGCTTAC	324	

## RESULT 11

E16378	E16378	345 bp	DNA	linear	PAT 28-JUL-1999
LOCUS	cDNA encoding heavy chain variable region of monoclonal antibody 2D3.				
DEFINITION	E16378				
ACCESSION	E16378				
VERSION	E16378.1 GI:5711061				
KEYWORDS	JP 1998165184-A/1.				
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	1 (bases 1 to 345)				
AUTHORS	Miyata, K., Anzai, M. and Shigeta, K.				
TITLE	ANTIBODY, GENE AND PRODUCTION OF CHIMERA ANTIBODY				
JOURNAL	Patent: JP 1998165184-A 1 23-JUN-1998.				

## COMMENT

PN	JP 1998165184-A/1	
PD	23-JUN-1998	
PF	16-DEC-1996	JP 1996335743
PI	MIYATA KENJI, ANZAI MAGAYUKI, SHIGETA KATSUMI	PC
CC	C12N15/09, C12P21/08, (C12P21/08, C12R1:91);	
CC	strandedness: Double;	
CC	topology: Linear;	
FH	Key	Location/Qualifiers

FEATURES  
SOURCE

```

1: 313
source
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

```

## ORIGIN

Query Match	69.7%	Score 23;	DB 6;	Length 345;
Best Local Similarity	83.9%;	Pred. No. 30;		
Matches	26: Conservative	0: Mismatches	5: Indels	0: Gaps

QY 3 CATTATTACGCTAGTTCCTGGTTGCTTAC 33  
||||||| || |||||  
Db 279 CATTATTACGTGTAACCTCCGGTTGCTTAC 303  
||||||| || |||||

## RESULT 12

AC125843/c	AC125843	245199 bp	DNA	linear	HTG 20-NOV-2002
LOCUS	Rattus norvegicus clone CH230-48E15, WORKING DRAFT SEQUENCE, 2				
DEFINITION	unordered pieces.				
ACCESSION	AC125843				
VERSION	GI:25138318				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLT0P.				
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Rattus norvegicus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				

REFERENCE  
AUTHORS  
1 (bases 1 to 246199)  
Muzny D.Marie., Metzker M.Lee., Abramson S., Adams C., Alder J., Allen C., Allen H., Albrooks S., Amin A., Anguiano D., Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H., Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F., Biwalo K., Blair J., Blankenburg K., Blyth P., Brown M., Bryant N., Buhay C., Burch P., Burrell K., Calderon E., Cardenas V., Carter K., Cavazos I., Ceasar H., Center A., Chacko J., Chavez D., Chen G., Chen Y., Chen X., Chu J., Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L., Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D., Delgado O., Denson S., Deramo C., Ding Y., Dinh H., Divya K., Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K., Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G., Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P., Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M., Gebregorgis E., Geer K., Gill R., Grady M., Guerra M., Guevara W., Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K., Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J., Hernandez R., Hines S., Hladun S.L., Hodgson A., Hognes M., Hollins B., Howells S., Hulyk S., Hume J., Idiebird D., Jackson A., Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A., Karpachy S., Kelly S., Kelly S., Khan Z., King L., Kovar C., Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J., Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J., Lorensuhewa L., Loulseghe H., Lozano R.J., Lu X., Ma J., Maheshwari M., Mahindratne M., Mahmoud M., Malloy K., Mangum A., Mangum B., Mapua P., Martin K., Martin R., Martinez E., Mawhiney S., McLeod M.P., McNeill T.Z., Meenen E., Milosavljevic A., Minet G., Minja E., Montemayor J., Moore S., Morgan M., Morris K., Morris S., Mundaesa M., Murphy M., Nair L., Nankervis C., Neal D., Newton N., Nguyen N., Norris S., Nwakeleneh O., Okwunigbo G., Olarnpungsoo A., Pal S., Parks K., Pasternak S., Paul H., Perez A., Perez L., Pfamkoch C., Plopper P., Poinexter A., Popovic D., Primus E., Pu L.-L., Puzo M., Quiroz J., Rachin E., Reeves K., Regier M.A., Reigh R., Reilly B., Reilly M., Ren Y., Reuter M., Richards S., Riggs F., Rives C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.J., Sanders W., Savary G., Scherer S., Scott G., Shatsman S., Shen H., Shetty J., Shvartsbeyn A., Sison I., Sitter C.D., Smajls D., Sneed A., Sodergren E., Song X.-Z., Sorelle R., Sosa J., Steimle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C., Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Usmani K., Valas R., Vera V., Villanasa D., Waldron L., Walker B., Wang J., Wang Q., Wang S., Villasana D., Waldron L., Walker B., White F., Williams G., Willson R., Wleczyk R., Wooden H., Worley K., Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V., Yu P., Zhang J., Zhou J., Zhou X., Zhao S., Dunn D., von Niederhausern A., Weiss R., Smith D.R., Holt R.A., Smith H.O., Weinstock G. and Gibbs R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 246199)  
Worley K.C.  
Direct Submission  
TITLE  
JOURNAL  
AUTHORS  
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 246199)  
Rat Genome Sequencing Consortium.  
Direct Submission  
TITLE  
JOURNAL  
AUTHORS  
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
COMMENT  
On Nov 20, 2002 this sequence version replaced gi:23195392.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GZDX  
Center clone name: CH230-48E15  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 218328 bases at least Q40  
Consensus quality: 220405 bases at least Q30  
Consensus quality: 221720 bases at least Q20  
Estimated insert size: 222748; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved  
\* \*  
\* 1 29617: contig of 29617 bp in length  
\* 29618 29717: gap of unknown length  
\* 29718 246199: contig of 216482 bp in length.  
FEATURES  
source  
1. 246199  
Location/Qualifiers  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clones="CH230-48E15"  
1. 1124  
misc\_feature  
/note="wgs end extension  
clone\_end:Sp6"  
1021..2029  
misc\_feature  
/note="clone boundary  
clone\_end:Sp6  
site:  
end sequence:BH363317"  
7743..8308  
misc\_feature  
/note="clone boundary  
clone\_end:T7  
site:  
end\_sequence:BH363315"  
ORIGIN  
Query Match 69.7%; Score 23; DB 2; Length 246199;  
Best Local Similarity 83.9%; Pred. No. 21;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 2 CCATTATTACGCTAGTTCCTGTTGCTTA 32  
||||||| ||||| ||||| ||||| |||||  
Db 245756 CCATTATTACCTAGTTCCTGTTGCTTA 245726  
RESULT 13  
LOCUS  
AC130129  
DEFINITION Rattus norvegicus clone CH230-69U12, WORKING DRAFT SEQUENCE, 4  
unordered pieces.  
ACCESSION  
AC130129  
VERSION  
AC130129.4 GI:30521324  
KEYWORDS  
HTG: HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
SOURCE  
Rattus norvegicus (Norway rat)



**AUTHORS**  
Shimoda, M., Nakamura, T., Takahashi, Y., Asanuma, H., Tamura, S., Kurata, T., Mizuuchi, T., Azuma, N., Kanno, C. and Takemori, T.  
**TITLE**  
Isotype-specific selection of high affinity memory B cells in nasal-associated lymphoid tissue  
**JOURNAL**  
J. Exp. Med. 194 (11), 1597-1607 (2001)  
**MEDLINE**  
21590495  
**PUBMED**  
11733574  
**REFERENCE**  
2 (bases 1 to 375)  
**AUTHORS**  
Shimoda, M.  
**Direct Submission**  
**TITLE**  
Submitted (19-MAR-2001) Michiko Shimoda, Utsunomiya University, Applied Biochemistry; 350 Mine-machi, Utsunomiya, Tochigi 321-8505, Japan (E-mail: ametan@cc.utsunomiya-u.ac.jp, Tel: 81-286-49-5478)  
**FEATURES**  
source  
1..375  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/isolate="PMG1510"  
/db\_xref="taxon:10090"  
/cell\_type="PCLN memory B cell"  
1..375  
/gene="V186.2-D-JH3-Cgamma2b"  
1..375  
/gene="V186.2-D-JH3-Cgamma2b"  
/codon\_start=1  
/product="immunoglobulin heavy chain"  
/protein\_id="BAB79133.1"  
/db\_xref="GI:17425999"  
/translation="QVQLQQPQAEVLKPGASVKLSCKASGYTFETYNMHWKQRPGRG LEWIGRIDPDSGGTKYNEKEKSKATLTVDKPSSTAFMQLSLTSDSAVYVCARYDYG SSWFAYWGQGLTIVTSAKYTPP"  
**ORIGIN**  
Query Match 69.1%; Score 22.8; DB 10; Length 375;  
Best Local Similarity 92.3%; Pred. No. 37;  
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
**Qy** 8 ATTACGCTAGTTCCTGGTTGCTTAC 33  
|||||  
**Db** 299 ACTACGGTAGTTCCTGGTTGCTTAC 324  
|||||  
**RESULT 15**  
AF072776 333 bp mRNA linear ROD 18-JUL-1998  
**LOCUS**  
Mus musculus 375.96 monoclonal anti-DNA IgM heavy chain variable region mRNA, partial cds.  
**DEFINITION**  
AF072776  
**ACCESSION**  
AF072776.1 GI:3328074  
**VERSION**  
AF072776.1  
**KEYWORDS**  
Mus musculus (house mouse)  
**SOURCE**  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**REFERENCE**  
1 (bases 1 to 333)  
**AUTHORS**  
Krishnan, M.R. and Marion, T.N.  
**TITLE**  
Comparison of the frequencies of arginines in heavy-chain CDR3 of antibodies expressed in the primary B cell repertoires of autoimmune-prone and normal mice  
**JOURNAL**  
Unpublished  
**REFERENCE**  
2 (bases 1 to 333)  
**AUTHORS**  
Krishnan, M.R. and Marion, T.N.  
**Direct Submission**  
**TITLE**  
Submitted (16-JUN-1998) Microbiology and Immunology, The University of Tennessee, Memphis, 858 Madison Ave., Memphis, TN 38163, USA  
**JOURNAL**  
Location/Qualifiers  
**FEATURES**  
source  
1..333  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="(NZB x NZW)F1"  
/db\_xref="taxon:10090"  
/cell\_line="375.96"

/cell\_type="B cell hybridoma"  
/rearranged  
/note="generated from LPS-stimulated spleen cells"  
<1..333  
/codon\_start=3  
/product="monoclonal anti-DNA IgM heavy chain variable region"  
/protein\_id="AAC26769.1"  
/db\_xref="GI:3328075"  
/translation="VOLKESGPGLVAPQSLSITCTVSGFSLTSYAI5VWVPPQKGL EWLGIVTGGTNTNSALKSRLSISKNSQVFLKWNLSLOTDDTARTYCAIYGVSSP FAYWGQGT"  
**ORIGIN**  
Query Match 67.9%; Score 22.4; DB 10; Length 333;  
Best Local Similarity 81.2%; Pred. No. 56;  
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
**Qy** 2 CCATTATTACGCTAGTTCCTGGTTGCTTAC 33  
|||||  
**Db** 286 CCATTATTACGCTAGTTCCTGGTTGCTTAC 317  
|||||  
**RESULT 16**  
AC143115 121360 bp DNA linear HTG 09-APR-2003  
**LOCUS**  
Macaca mulatta clone CH250-269J6, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
**DEFINITION**  
AC143115  
**ACCESSION**  
AC143115.1 GI:29567754  
**VERSION**  
HTG: HTGS PHASE2; HTGS PGI.  
**KEYWORDS**  
Macaca mulatta (rhesus monkey)  
**SOURCE**  
Macaca mulatta  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.  
**REFERENCE**  
1 (bases 1 to 121360)  
**AUTHORS**  
Cueros, M. and Milosavljevic, A.  
**TITLE**  
Pooled genomic indexing (PGI): mathematical analysis and experiment design  
**JOURNAL**  
(in) Guigo, R. and Gusfield, D. (Eds.);  
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI 2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28; Springer (2002)  
2 (bases 1 to 121360)  
**REFERENCE**  
Milosavljevic, A., Sodergren, E., Cueros, M., Li, B., Jackson, A.R., Adams, C., Adio-Oduola, B., Ali-ouman, F.R., Allen, C., Alsbrooks, S.L., Amarant, J., Bim, K., Ayele, M., Banks, T., Barbata, J., Benton, J., Bim, K., Blau, K., Brown, M., Bryant, N.P., Buhay, C., Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, H.H., Delaney, K.R., Delgado, O., Degen, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Egan, A., Earnhart, C., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H., Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzger, M., Miller, A., Miner, G., Miner, G., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N.,

```

Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Oguh,M., Okwundu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
Shoostari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A.,
Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Watlington,S., Williams,G., Williamson,A., Zhou,J.,
Wieczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
Zorrilla,S., Kucheralapati,R., Weinstock,G. and Gibbs,R.
Direct Submission
3 (bases 1 to 121360)
Worley,K.C.
Direct Submission
Submitted (05-APR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 121360)
Worley,K.C.
Direct Submission
Submitted (09-APR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: LBDF
Center clone name: CH250-26936
----- Summary Statistics
Chemistry: Dye-primer Bodypy: inf% of reads
Chemistry: Dye-terminator Big Dye: inf% of reads
Consensus quality: 3836 bases at least Q40
Consensus quality: 4687 bases at least Q30
Consensus quality: 5697 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: The contigs are based on the application
* of the PGI method using the Human genome (NCBI build 31)
* as the comparative genome.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 121360: contig of 121360 bp in length.
FEATURES
source
1..121360
/organism="Macaca mulatta"
/mol_type="genomic DNA"
/db_xref="taxon:9544"
/clone="CH250-269J6"
misc_feature
1..121360
/notes="assembly name:CH250-269J6.1A
CONFIDENCE:0.67"
ORIGIN
Query Match 66.7%; Score 22; DB 2; Length 121360;
Best Local Similarity 83.3%; Pred. No. 60;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CATTATTACGCTAGTTCCTGGTTGCTTA 32

```

```

Db 104918 CATTATTTCCTAATTCCTGAGTTGCTTA 104947
||||| ||||| ||||| ||||| |||||
AC132625 183472 bp DNA linear HTG 15-SEP-2004
Mus musculus chromosome 1 clone RP23-101M7, WORKING DRAFT SEQUENCE,
5 unordered pieces.
AC132625 GI:52078003
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ACTIVEFIN.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Wilson,R.K.
1 (bases 1 to 183472)
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 183472)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (01-SEP-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 183472)
Wilson,R.K.
Direct Submission
Submitted (15-SEP-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Sep 15, 2004 this sequence version replaced gi:22597729.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_BA0101M07
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 182146 bases at least Q40
Consensus quality: 182337 bases at least Q30
Consensus quality: 182486 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 16533: contig of 16533 bp in length
* 16534 16633: gap of unknown length
* 16634 29068: contig of 12435 bp in length
* 29069 52824: gap of unknown length
* 52825 52924: contig of 23656 bp in length
* 52925 74398: gap of unknown length
* 74399 74499: contig of 21474 bp in length
* 74499 183472: contig of 108974 bp in length.
FEATURES
source
1..183472
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"

```



in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GJRH  
Center clone name: CH230-230D24  
----- Summary Statistics  
Assembly program: Atlas 3.0;  
Consensus quality: 239462 bases at least Q40  
Consensus quality: 242145 bases at least Q30  
Estimated insert size: 236442 bases at least Q20  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html)  
\* NOTE: This sequence may represent more than one clone.  
\* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

\* 1 246951: contig of 246951 bp in length.

Location/Qualifiers  
1..246951  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-230D24"

misc\_feature  
1..1288  
/note="wgs end extension  
clone\_end:Sp6"  
3146...3989  
clone\_end:Sp6  
clone\_end:Sp6 boundary  
site:ECORI  
end sequence:BZ111913"  
complement(240540..241289)  
/note="clone boundary  
clone\_end:T7"  
site:ECORI  
end sequence:RWBOH24TJ"  
244717..246951  
/note="wgs end extension  
clone\_end:T7"

misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature

ORIGIN  
Query Match 66.1%; Score 21.8; DB 2; Length 246951;  
Best Local Similarity 78.8%; Pred. NO. 70;  
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCATTTATTACGGTAGTTCCTCGGTTCGTTTC 33  
Db 193346 CCATTTTAACCTAGTTCAGGGTATCTTC 193314

RESULT 21  
MUSIGDAJ



```

LOCUS      MUSIGDJAJ              37 bp      DNA      linear      ROD 02-AUG-1995
DEFINITION Mus musculus rearranged immunoglobulin heavy chain D-J junction
ACCESSION  L36732
VERSION    L36732.1 GI:927126
KEYWORDS   diversity region; immunoglobulin heavy chain; joining region;
           rearranged.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS    Chowers, Y., Harwood, J., Holtmeier, W., Kagnoff, M.P. and
           Morzycka-Wroblewska, B.
TITLE      No title
JOURNAL    Unpublished (1994)
COMMENT    Original source text: Mus musculus (clone: 40JH3) bone marrow DNA.
FEATURES   Location/Qualifiers
           1..37
             /organism="Mus musculus"
             /mol_type="genomic DNA"
             /db_xref="taxon:10090"
             /clone="40JH3"
             /tissue_type="bone marrow"
           misc_feature 1..37
             /note="D-J junction"
ORIGIN
Query Match      65.5%; Score 21.6; DB 10; Length 37;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy  6 TTATTACGCTAGTTCCTGGTTTGCTTAC 33
    ||| ||||| ||||| ||||| ||||| |||||
Db  5 TTACTACGGTAGAGCCTGGTTTGCTTAC 32

RESULT 22
AY182575
LOCUS      Mus musculus clone L2NL-PQ7 immunoglobulin heavy chain variable
DEFINITION region mRNA, partial cds.
ACCESSION  AY182575
VERSION    AY182575.1 GI:27728941
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Mus musculus
REFERENCE  1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS    Kretschmer, K., Engel, H. and Weiss, S.
TITLE      Strong antigenic selection shaping the immunoglobulin heavy chain
           repertoire of B-1a lymphocytes in lambda 2(315) transgenic mice
JOURNAL    Eur. J. Immunol. 32 (8), 2317-2327 (2002)
MEDLINE    22197943
PubMed     12209645
REFERENCE  2 (bases 1 to 162)
AUTHORS    Kretschmer, K., Engel, H. and Weiss, S.
TITLE      Direct Submission
JOURNAL    Submitted (18-NOV-2002) Molecular Immunology, GBF, German Research
           Centre for Biotechnology, Mascheroder Weg 1, Braunschweig 38124,
           Germany
FEATURES   Location/Qualifiers
           1..162
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="BALB/c"
             /db_xref="taxon:10090"
             /clone="L2NL-PQ7"
             /cell_type="B lymphocyte"
             /tissue_type="liver"
             /dev_stage="neonate"
             /note="lambda2 MOPC315 transgenic L2 mouse"
ORIGIN
Query Match      65.5%; Score 21.6; DB 10; Length 330;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy  6 TTATTACGCTAGTTCCTGGTTTGCTTAC 33
    ||| ||||| ||||| ||||| ||||| |||||
Db  270 TTACTACGGTAGTAGCTGGTTTGCTTAC 297

```

```

CDS
<1..>162
/codon_start=1
/product="immunoglobulin heavy chain variable region"
/protein_id="AA018931.1"
/db_xref="GI:27728942"
/tranlation="RLSISKDNKSQVFLKNSLQTDITAMYCARDHYGSSWFAWM
GQGLTVTVSA"
ORIGIN
Query Match      65.5%; Score 21.6; DB 10; Length 162;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy  6 TTATTACGCTAGTTCCTGGTTTGCTTAC 33
    ||| ||||| ||||| ||||| ||||| |||||
Db  102 TTACTACGGTAGTAGCTGGTTTGCTTAC 129

RESULT 23
AY182599
LOCUS      Mus musculus clone L2FL-P1 immunoglobulin heavy chain variable
DEFINITION region mRNA, partial cds.
ACCESSION  AY182599
VERSION    AY182599.1 GI:27728989
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Mus musculus
REFERENCE  1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS    Kretschmer, K., Engel, H. and Weiss, S.
TITLE      Strong antigenic selection shaping the immunoglobulin heavy chain
           repertoire of B-1a lymphocytes in lambda 2(315) transgenic mice
JOURNAL    Eur. J. Immunol. 32 (8), 2317-2327 (2002)
MEDLINE    22197943
PubMed     12209645
REFERENCE  2 (bases 1 to 330)
AUTHORS    Kretschmer, K., Engel, H. and Weiss, S.
TITLE      Direct Submission
JOURNAL    Submitted (18-NOV-2002) Molecular Immunology, GBF, German Research
           Centre for Biotechnology, Mascheroder Weg 1, Braunschweig 38124,
           Germany
FEATURES   Location/Qualifiers
           1..330
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="BALB/c"
             /db_xref="taxon:10090"
             /clone="L2FL-P1"
             /cell_type="B lymphocyte"
             /tissue_type="liver"
             /dev_stage="fetus"
             /note="lambda2 MOPC315 transgenic L2 mouse"
ORIGIN
Query Match      65.5%; Score 21.6; DB 10; Length 330;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy  6 TTATTACGCTAGTTCCTGGTTTGCTTAC 33
    ||| ||||| ||||| ||||| ||||| |||||
Db  270 TTACTACGGTAGTAGCTGGTTTGCTTAC 297

```

```

RESULT 24
CR478208      198284 bp      DNA      linear      HTG 11-JUN-2004
LOCUS
DEFINITION    Danio rerio clone CH211-149M3, WORKING DRAFT SEQUENCE, 8 unordered
               pieces
ACCESSION
CR478208      GI:48639784
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE        Danio rerio (zebrafish)
ORGANISM      Danio rerio
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
               Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 198284)
   Burton, J.
   Direct Submission
   Submitted (10-JUN-2004) Wellcome Trust Sanger Institute, Hinxton,
   Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
   zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
   ----- Genome Center
   Center: Wellcome Trust Sanger Institute
   Center code: SC
   Web site: http://www.sanger.ac.uk
   Contact: zfsh-help@sanger.ac.uk
   ----- Project Information
   Center project name: ZC149M3
   ----- Summary Statistics
   Assembly program: XGAP4; version 4.5
   Chemistry: Dye-terminator; 100% of reads
   Consensus quality: 195080 bases at least Q40
   Consensus quality: 195650 bases at least Q30
   Consensus quality: 196011 bases at least Q20
   Insert size: 197584; sum-of-contigs
   Insert size: 202177; 1.7% error; agarose-fp
   Quality coverage: 7.19x in Q20 bases; sum-of-contigs Quality
   Coverage: 7.16x in Q20 bases; agarose-fp
   -----
   * NOTE: This is a 'working draft' sequence. It currently
   * consists of 8 contigs. The true order of the pieces
   * is not known and their order in this sequence record is
   * arbitrary. Gaps between the contigs are represented as
   * runs of N, but the exact sizes of the gaps are unknown.
   * This record will be updated with the finished sequence
   * as soon as it is available and the accession number will
   * be preserved.
   *
   1
   4861: contig of 4861 bp in length
   4862
   4961: gap of 100 bp
   4962
   65190: contig of 60229 bp in length
   65191
   83695: gap of 100 bp
   83696
   83795: contig of 18405 bp in length
   83796
   111212: contig of 100 bp
   111213
   111212: gap of 100 bp
   111213
   120577: contig of 9265 bp in length
   120578
   120677: gap of 100 bp
   120678
   163275: contig of 42598 bp in length
   163276
   163376: gap of 100 bp
   163376
   170139: contig of 6764 bp in length
   170140
   170239: gap of 100 bp
   170240
   198284: contig of 28045 bp in length.
   Location/Qualifiers
   1. .198284
   /organism="Danio rerio"
   /mol_type="genomic DNA"
   /db_xref="taxon:7955"
   /clone="CH211-149M3"
   /clone_lib="CHORI-211"
   /notes="assembly fragment:00016
   fragment chain:1
   clone end:T7
   vector_side:left"
   misc_feature
   1. .4861
   /note="assembly fragment:00016
   fragment chain:1
   clone end:T7
   vector_side:left"
   misc_feature
   4962. .65190
   /note="assembly fragment:01429
   fragment chain:1"
   65291. .83695
   /note="assembly fragment:00231
   fragment chain:1"
   83796. .111212
   /note="assembly fragment:00724
   fragment chain:2"
   111313. .120577
   /note="assembly fragment:00141
   fragment chain:2"
   120678. .163275
   /note="assembly fragment:01017
   fragment chain:2"
   163376. .170139
   /note="assembly fragment:00078
   fragment chain:2"
   170240. .198284
   /note="assembly fragment:00436
   clone end:SP6
   vector_side:right"
ORIGIN
Query Match      65.5%; Score 21.6; DB 2; Length 198284;
Best Local Similarity 85.7%; Pred. No. 87;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 TTATTACGCTAGTTCCTGCTTTCCTTA 32
   |||||
Db 73422 TTTATTACGCTAGTTCCTGCTTTCCTTA 73449

RESULT 25
AC018369      161671 bp      DNA      linear      HTG 26-MAR-2000
LOCUS
DEFINITION    Homo sapiens chromosome 3 clone RP11-274H2 map 3, WORKING DRAFT
               SEQUENCE, 12 unordered pieces.
ACCESSION
AC018369      GI:7329294
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE        Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 161671)
   Birren, B., Linton, L., Nusbaum, C., Lander, E., Abramson, H., Allen, N.,
   Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
   Boguslavsky, L., Bouckgaier, B., Brown, A., Castle, A., Colangelo, M.,
   Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K.,
   Domino, M., Doyle, M., Fenestor, J., Ferreira, P., FitzHugh, W.,
   Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B.,
   Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L.,
   Karatas, A., Klein, J., Landers, T., Lehoczy, J., Lieu, C., Locke, K.,
   Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
   Meldrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T.,
   O'Donnell, P., Peterson, K., Pierre, N., Pollara, V., Riley, R.,
   Rothman, D., Roy, A., Santos, R., Severy, P., Stange-Thomann, N.,
   Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
   Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
   Zimmer, A. and Zody, M.
   Direct Submission
   Submitted (10-DEC-1999) Whitehead Institute/MIT Center for Genome
   Research, 320 Charles Street, Cambridge, MA 02141, USA
   On Mar 26, 2000 this sequence version replaced gi:6642724.
   All repeats were identified using RepeatMasker:
   Smit, A.P.A. & Green, P. (1996-1997)
   http://ftp.genome.washington.edu/RM/RepeatMasker.html
   ----- Genome Center
   Center: Whitehead Institute/ MIT Center for Genome Research

```

Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L1811  
 Center clone name: 274 H.2  
 ----- Summary Statistics  
 Sequencing vector: M13; M7815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 155036 bases at least Q40  
 Consensus quality: 157760 bases at least Q30  
 Consensus quality: 158947 bases at least Q20  
 Insert size: 163000; agarose-fp  
 Insert size: 160571; sum-of-contigs  
 Quality coverage: 5.0 in Q20 bases; agarose-fp  
 Quality coverage: 5.1 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 1285: contig of 1285 bp in length  
 \* 1286 1385: gap of 100 bp  
 \* 1386 2871: contig of 1486 bp in length  
 \* 2872 2971: gap of 100 bp  
 \* 2972 4246: contig of 1275 bp in length  
 \* 4247 4346: gap of 100 bp  
 \* 4347 8560: contig of 4214 bp in length  
 \* 8561 8661: gap of 100 bp  
 \* 8662 16872: contig of 8212 bp in length  
 \* 16873 25499: contig of 8527 bp in length  
 \* 25500 25599: gap of 100 bp  
 \* 25600 36640: contig of 11041 bp in length  
 \* 36641 37640: gap of 100 bp  
 \* 37641 47799: contig of 11059 bp in length  
 \* 47800 47899: gap of 100 bp  
 \* 47900 64902: contig of 17003 bp in length  
 \* 64903 65002: gap of 100 bp  
 \* 65003 79335: contig of 14333 bp in length  
 \* 79336 79435: gap of 100 bp  
 \* 79436 112959: contig of 33524 bp in length  
 \* 112960 113059: gap of 100 bp  
 \* 113060 161671: contig of 48612 bp in length.  
 ----- Location/Qualifiers  
 1. 161671  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /map="3"  
 /clone="RP11-274H2"  
 /clone\_lib="RPCI-11 Human Male BAC"  
 1. 1285  
 /note="assembly\_fragment"  
 1386. 2871  
 /note="assembly\_fragment"  
 2972. 4246  
 /note="assembly\_fragment"  
 4347. 8560  
 /note="assembly\_fragment"  
 clone\_end.SP6  
 vector\_side:right  
 8661. 16872  
 /note="assembly\_fragment"  
 16973. 25499  
 /note="assembly\_fragment"  
 clone\_end.T7

vector\_side:right  
 25600..36640  
 /note="assembly\_fragment"  
 36741..47799  
 /note="assembly\_fragment"  
 47900..64902  
 /note="assembly\_fragment"  
 65003..79335  
 /note="assembly\_fragment"  
 79436..112959  
 /note="assembly\_fragment"  
 113060..161671  
 /note="assembly\_fragment"  
 ORIGIN  
 Query Match 64.8%; Score 21.4; DB 2; Length 161671;  
 Best Local Similarity 80.6%; Pred. No. 1.le+02;  
 Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1 CCATTTATTACGCTAGTTCCTGGTTGCTT 31  
 |||||  
 Db 125929 CCCATTTAAACCTATTTTCATTTGCTT 125959  
 RESULT 26  
 AC107021/c  
 LOCUS  
 DEFINITION Homo sapiens 3 BAC RP11-274H2 (Roswell Park Cancer Institute Human  
 BAC Library) complete sequence.  
 AC107021  
 AC107021.4 GI:19033423  
 HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 161731)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
 Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,  
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,  
 Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,  
 Jackson,L.E., Jacobson,B., Jia,X., Johnson,R., Jolivet,S.,  
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,  
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
 Loulseg,H., Lozado,R., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,  
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,  
 Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,  
 Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,  
 Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,  
 Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,  
 Oriedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,  
 Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,  
 Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,  
 Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,B.,  
 Sonaika,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,  
 Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,  
 Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vaquez,L.,

Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S.,  
 Warren, R., Washington, C., Watlington, S., Williams, G.,  
 Williamson, A., Wlezyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,  
 Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstock, G. and  
 Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 161731)  
 Worley, K.C.  
 Direct Submission  
 Submitted (14-JAN-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 161731)  
 Worley, K.C.  
 Direct Submission  
 Submitted (22-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 4 (bases 1 to 161731)  
 Worley, K.C.  
 Direct Submission  
 Submitted (01-MAR-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Mar 1, 2002 this sequence version replaced gi:18854923.  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

## COMMENT

CLONE LENGTH: This sequence does not necessarily represent the  
 entire insert of this clone. Overlapping regions of clones are only  
 sequenced and submitted once, so the sequence for the remainder of  
 the insert may be found in the record for the adjacent clones.  
 Overlapping clones are noted at the beginning and end of the  
 Features listing.

## ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches  
 of a local database that includes entries from dbSTS, GDB, and  
 local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,  
 unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST  
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
 EST and cDNA sequences. Genes demonstrate at least two exons  
 flanked by consensus splice sites that maintained sequence  
 continuity across the splice junctions. Sequences that are not  
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
 standard of double strand coverage with a minimum of 2 clones and 2  
 reads with no ambiguities or 2 chemistries with a minimum of 2  
 clones and 3 reads with no ambiguities. If the sequence quality for  
 a region does not meet this standard, it will be indicated in the  
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
 standards - estimated error rate less than 1 per 10,000 bases.  
 Reports of lowest quality individual bases and measures of base  
 quality are listed below. Description of the metrics can be found  
 at URL:  
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT.

FEATURES  
 source  
 Location/Qualifiers  
 1. .161731  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /clone="RP11-274H2"  
 1. .754  
 /rpt\_family="L1PA16"  
 repeat\_region

repeat\_region  
 complement(755..1042)  
 /rpt\_family="AluSx"  
 1043..1637  
 /rpt\_family="L1PA16"  
 1644..2024  
 /rpt\_family="MSTB"  
 2025..3254  
 /rpt\_family="L1PA16"  
 3284..3318  
 /rpt\_family="AT\_rich"  
 3343..3495  
 /rpt\_family="(TTATA)n"  
 3726..3959  
 /rpt\_family="L2"  
 3965..4004  
 /rpt\_family="(TGAA)n"  
 4314..4353  
 /rpt\_family="MSTD"  
 4674..4698  
 /rpt\_family="AT\_rich"  
 5022..5059  
 /rpt\_family="(TG)n"  
 complement(5156..5241)  
 /rpt\_family="MLT2CB"  
 complement(5236..5324)  
 /rpt\_family="MLT2G"  
 complement(5325..5753)  
 /rpt\_family="MLT2CB"  
 5789..6390  
 /rpt\_family="LTR19B"  
 complement(6391..6690)  
 /rpt\_family="AluY"  
 complement(6804..6948)  
 /rpt\_family="L2"  
 complement(7155..7489)  
 /rpt\_family="MER2"  
 7491..7634  
 /rpt\_family="(TA)n"  
 complement(7931..8523)  
 /rpt\_family="MLT1E2"  
 9037..9171  
 /rpt\_family="MER5A"  
 9171..9343  
 /rpt\_family="MER5A"  
 9353..9522  
 /rpt\_family="MER5B"  
 complement(9558..9643)  
 /rpt\_family="MER5B"  
 complement(10017..10190)  
 /rpt\_family="L1MCC"  
 10200..10257  
 /rpt\_family="LTR24"  
 10451..10665  
 /rpt\_family="LTR24B"  
 10717..11076  
 /rpt\_family="Tigger2a"  
 complement(11093..11238)  
 /rpt\_family="L1MCC"  
 11239..12066  
 /rpt\_family="MER21B"  
 complement(12067..12396)  
 /rpt\_family="L1MCC"  
 complement(12569..13013)  
 /rpt\_family="L1MCC"  
 13015..13078  
 /rpt\_family="MER5B"  
 13820..14132  
 /rpt\_family="AluSx"  
 14463..14507  
 /rpt\_family="(CA)n"  
 15247..15269  
 /rpt\_family="AT\_rich"  
 complement(15589..15683)

```

repeat_region      /rpt_family="MER63"
17012..17036
/rpt_family="AT_rich"
repeat_region      17492..17892
/rpt_family="L1M4"
repeat_region      17978..18208
/rpt_family="L1MB2"

Query Match      64.8%; Score 21.4; DB 9; Length 161731;
Best Local Similarity 80.6%; Pred. No. 1.1e+02;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCCATTATTACGCTAGTCTCTGGTTGGCTT 31
Db 104920 CCCATTAAACTCTATTTCATTGTTGCTT 104890

RESULT 27
AC150560          225551 bp      DNA      linear      HTG 06-AUG-2004
LOCUS             Mus musculus chromosome 1 clone RP23-16114, WORKING DRAFT
DEFINITION        SEQUENCE, 14 unordered pieces.
ACCESSION          AC150560
VERSION            AC150560.2 GI:51011350
KEYWORDS           HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE             Mus musculus (house mouse)
ORGANISM           Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 225551)
Wilson,R.K.
Unpublished
2 (bases 1 to 225551)
Wilson,R.K.
Direct Submission
Submitted (24-JUL-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 225551)
Wilson,R.K.
Direct Submission
Submitted (06-AUG-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Aug 6, 2004 this sequence version replaced gi:50582874.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M BA0016114
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 217350 bases at least Q40
Consensus quality: 218582 bases at least Q30
Consensus quality: 219211 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1100: contig of 1100 bp in length
1101 1200: gap of unknown length

repeat_region      /rpt_family="MER63"
17012..17036
/rpt_family="AT_rich"
repeat_region      17492..17892
/rpt_family="L1M4"
repeat_region      17978..18208
/rpt_family="L1MB2"

Query Match      64.8%; Score 21.4; DB 2; Length 225551;
Best Local Similarity 80.6%; Pred. No. 1.1e+02;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 CCATTATTACGCTAGTCTCTGGTTGGCTTA 32
Db 201620 CTATCCATTAAAGCTAGTTCACGGTTTGATTA 201650

FEATURES
            Location/Qualifiers
                source
                    1..225551
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10090"
                        /chromosome="15"
                        /clone="RP23-16114"
                misc_feature
                    1..1100
                        /notes="assembly_name:Contig51"
                misc_feature
                    1201..3320
                        /note="assembly_name:Contig52"
                        clone_end:T7
                        vector_side:left
                    3421..6288
                        /notes="assembly_name:Contig54"
                    6389..10343
                        /notes="assembly_name:Contig55"
                    10444..16673
                        /notes="assembly_name:Contig56"
                    16774..22742
                        /notes="assembly_name:Contig57"
                    22843..33670
                        /notes="assembly_name:Contig58"
                    33771..44715
                        /notes="assembly_name:Contig59"
                    44816..54386
                        /notes="assembly_name:Contig60"
                    54487..66689
                        /notes="assembly_name:Contig61"
                    66790..91617
                        /notes="assembly_name:Contig62"
                    91718..110757
                        /notes="assembly_name:Contig63"
                    110858..147945
                        /notes="assembly_name:Contig64"
                    148046..225551
                        /notes="assembly_name:Contig65"

ORIGIN
Query Match      64.8%; Score 21.4; DB 2; Length 225551;
Best Local Similarity 80.6%; Pred. No. 1.1e+02;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 CCATTATTACGCTAGTCTCTGGTTGGCTTA 32
Db 201620 CTATCCATTAAAGCTAGTTCACGGTTTGATTA 201650

```



AC108553	AC108553.4	GI:30581394	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.	HTGS norvegicus (Norway rat)	Rattus norvegicus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1	(bases 1 to 250348)	Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alibrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, P., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyte, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, T., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowls, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Loulseged, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeleme, O., Okwuon, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojars, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shivartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.			
REFERENCE	2	(bases 1 to 250348)	Unpublished	Worley, K.C.	Direct Submission	Submitted (31-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3	(bases 1 to 250348)	Submitted (13-MAY-2003)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	Direct Submission	Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	4	(bases 1 to 250348)	Submitted (15-NOV-2002)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	Direct Submission	Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

#### REFERENCE AUTHORS

1 (bases 1 to 272952)  
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooke, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Loulseghe, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Manthey, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naik, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, I., Rojars, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmami, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, J., Zhao, J., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.  
Direct Submission  
Unpublished  
2 (bases 1 to 272952)  
Worley, K. C.  
Direct Submission  
Submitted (07-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 272952)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 15, 2002 this sequence version replaced gi:23266116.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GPVI  
Center clone name: CH230-78J24  
----- Summary Statistics

Assembly program: Phrap; version 0.990329  
Consensus quality: 210764 bases at least Q40  
Consensus quality: 215490 bases at least Q30  
Consensus quality: 218470 bases at least Q20  
Estimated insert size: 212334; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 272952: contig of 272952 bp in length.

#### FEATURES

source  
1. .272952  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-78J24"  
18757. .19959  
/note="wgs\_end\_extension  
clone\_end:Sp6"  
complement(19366. .19855)  
/note="clone\_boundary  
clone\_end:Sp6  
site:  
end\_sequence:BH292093"  
41635. .44174  
/note="wgs\_contig"  
51031. .52203  
/note="wgs\_contig"  
186330. .187412  
/note="wgs\_contig"  
195675. .197696  
/note="wgs\_contig"  
271499. .272310  
/note="clone\_boundary  
clone\_end:T7  
site:  
end\_sequence:BH292092"

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

#### ORIGIN

Query Match 64.2%; Score 21.2; DB 2; Length 272952;  
Best Local Similarity 88.5%; Pred. No. 1.3e+02;  
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 3 CATTATTACGCTAGTTCCTCGTTTG 28  
DB 41132 CTTTATTAGGTAGTTCCTCGTTTG 41157



```

RESULT 31
LOCUS      AF071679
DEFINITION Mus musculus clone MPJ-26 immunoglobulin heavy chain D-J region
            (VHJ558-D-J) gene, partial sequence.
ACCESSION  AF071679
VERSION     AF071679.1 GI:3320536
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE   1 (bases 1 to 50)
AUTHORS     Klonowski,K.D., Primiano,L.L. and Monestier,M.
TITLE       Atypical VH-D-JH rearrangements in newborn autoimmune MRL mice
JOURNAL     J. Immunol. 162 (3), 1566-1572 (1999)
MEDLINE     99138837
PUBMED      9973414
REFERENCE   2 (bases 1 to 50)
AUTHORS     Monestier,M. and Klonowski,K.
TITLE       Direct Submission
JOURNAL     Submitted (12-JUN-1998) Microbiology and Immunology, Temple
            University School of Medicine, 3400 N. Broad St., Philadelphia, PA
            19140, USA
FEATURES    Location/Qualifiers
             source          1..50
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /strain="MRL +/+"
                        /db_xref="taxon:10090"
                        /clone="MPJ-26"
                        /tissue_type="liver"
                        /dev_stage="newborn"
                        /rearranged
                        <1..>50
             gene            /gene="VHJ558-D-J"
             misc_feature    <1..>50
                        /gene="VHJ558-D-J"
                        /note="possible frameshift at D-J junction during
                        rearrangement may result in nonfunctional immunoglobulin
                        heavy chain"
ORIGIN
Query Match      63.6%; Score 21; DB 10; Length 50;
Best Local Similarity 82.8%; Pred. No. 2.6e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY  5  TTTATTACGGTAGTTCCTGGTTGCTTAC 33
    |||||  |||||  |||||  |||||  |||||
Db   2  TCTATTACTGCTGCTGCTGCTTAC 30

RESULT 32
LOCUS      AE008602
DEFINITION Rickettsia conorii str. Malish 7, section 34 of 114 of the complete
            genome.
ACCESSION  AE008602
VERSION     AE008602.1 GI:15619434
KEYWORDS
SOURCE      Rickettsia conorii str. Malish 7
            Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
            Rickettsiaceae; Rickettsiae; Rickettsia; spotted fever group.
            1 (sites)
REFERENCE   Oqata,H., Audic,S., Barbe,V., Artiguenave,F., Fournier,P.E.,
            Raoult,D., and Claverie,J.M.
TITLE       Selfish DNA in protein-coding genes of Rickettsia
JOURNAL     Science 290 (5490), 347-350 (2000)
MEDLINE     20485642
PUBMED      11030655
REFERENCE   2 (bases 1 to 13078)
AUTHORS     Oqata,H., Audic,S. and Claverie,J.-M.

```

```

TITLE       Selfish DNA and the origin of genes
JOURNAL     Science 291 (5502), 252-253 (2001)
REMARK      Response to Letter to the Editor
REFERENCE   3 (bases 1 to 13078)
AUTHORS     Oqata,H., Audic,S., Renesto-Audiffren,P., Fournier,P.-E., Barbe,V.,
            Samson,D., Roux,V., Cossart,P., Weissenbach,J., Claverie,J.-M. and
            Raoult,D.
TITLE       Mechanisms of evolution in Rickettsia conorii and R. prowazekii
JOURNAL     Science 293 (5537), 2093-2098 (2001)
MEDLINE     21442074
PUBMED      11557893
REFERENCE   4 (bases 1 to 13078)
AUTHORS     Oqata,H., Audic,S., Renesto-Audiffren,P., Fournier,P.-E., Barbe,V.,
            Samson,D., Roux,V., Cossart,P., Weissenbach,J., Claverie,J.-M. and
            Raoult,D.
TITLE       Direct Submission
JOURNAL     Submitted (26-APR-2001) Unite des Rickettsies, Faculte de Medecine,
            CNRS UMR 6020, Universite de la Mediteranee, 27 boulevard Jean
            Moulin, Marseille Cedex 05 13385, France
COMMENT     A public version of R. conorii genome database is accessible at
            http://igs-server.cnrs-mrs.fr/. The database intends to provide
            updated data. Annotation of the genome is an ongoing task whose
            goal is to make the genome sequence more useful. Comments to the
            authors are appreciated.
FEATURES    Location/Qualifiers
             source          1..13078
                        /organism="Rickettsia conorii str. Malish 7"
                        /mol_type="genomic DNA"
                        /strain="Malish 7"
                        /db_xref="taxon:272944"
                        910..1848
                        /locus_tag="RC0370"
                        910..1848
                        /locus_tag="RC0370"
                        /codon_start=1
                        /transl_table=11
                        /product="similarity to ATP-dependent protease subunit C
                        (ClpC)"
             gene            /protein_id="AAL02908.1"
                        /db_xref="GI:15619435"
                        /translation="MPFEDYCRKGLAHVPKSTSVSPAETENIFKAITDEAKIREWFD
                        KNYQVOEKIDNIAKSLNEENKFAILRRNLTNKSTSEIKNIVEKINMQVDIGKVL
                        GSAIGIKTLMHYLSYKWKGEKLNKFNVDVFRIRLKEILASWTYVYGTNIDELTSCF
                        VHYCLSDNLIKLEIDIKSIQDKRILLILLDGYDEVAFLSQSNRRDRDMDSVFQYKNIV
                        MSRPNAVIEENSRFEQKQDNTGWDSEGLEQVHKRNFYDKELGTPLKIFLDTHTGQI
                        KETCAVPNATIEICLVMSQDAIRDKFKTKTVIKILI"
                        2256..2645
                        /locus_tag="RC0371"
                        2256..2645
                        /locus_tag="RC0371"
                        /codon_start=1
                        /transl_table=11
                        /product="unknown"
                        /protein_id="AAL02909.1"
                        /db_xref="GI:15619436"
                        /translation="MTLKLPLAGVSNENNOELIEIFWKAVTCNVDRLLEIGIERKILL
                        LMLHLAQSNINGKFSRIPLNKLQIQLNLDIVLKDITGWEOHIDSGYLSAIVKTVN
                        EKLQNKTDTPQEFQVIGITGLANKK"
                        3361..3909
                        /locus_tag="RC0372"
                        3361..3909
                        /locus_tag="RC0372"
                        /codon_start=1
                        /transl_table=11
                        /product="unknown"
                        /protein_id="AAL02910.1"
                        /db_xref="GI:15619437"
                        /translation="MKEVLNDSGNEVKIVVLSLTVRINPSLAQETLKILNTLLNN
                        PSNYIFTAKILGWIIQINPNISHDASKILNLFNSDKSKSLSVLEGKVPVE
                        AFKVFYDILSDPYDVKYAIYAVIININVRFSFKASYKQVNRLLIKIIDLHRTQNP
                        DKDLXLEIDAKTTLQITIDYIT"
                        4823..5371
                        /locus_tag="RC0373"

```





```

Best Local Similarity 82.8%; Pred. No. 1.6e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 CCATTTATTACGCTAGTTCCTCGTTTCT 30
||||||| | | | | | | | | | | | | | |
Db 13859 CCATTTATTGCTGCTGTTCTCGTTTCT 13887

RESULT 36
AC108290 Rattus norvegicus clone CH230-201C19, WORKING DRAFT SEQUENCE, 5
LOCUS unorderd pieces.
DEFINITION AC108290
ACCESSION AC108290.5 GI:24818599
VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 223642)
Muzny,D.,Marle., Metzker,M.,Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyatebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,B.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flegg,N., Forbes,D., Foster,M., Foster,P.,
Fraser,C.M., Gabisl,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,T., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Haves,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,H., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenauhewa,L., Loulsegod,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwunu,G., Olarnpunseagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plommer,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmami,K.,
Valas,R., Vera,V., Villanasa,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished

REFERENCE
AUTHORS
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOVZ
Center clone name: CH230-201C19
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 198355 bases at least Q40
Consensus quality: 200628 bases at least Q30
Consensus quality: 202141 bases at least Q20
Estimated insert size: 200365; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 17684: contig of 17684 bp in length
* 17685 17784: gap of unknown length
* 17785 109488: contig of 91704 bp in length
* 109489 109588: gap of unknown length
* 109589 218503: contig of 108915 bp in length
* 218504 218603: gap of unknown length
* 218604 219707: contig of 1104 bp in length
* 219708 219807: gap of unknown length
* 219808 223642: contig of 3835 bp in length.
FEATURES
source
1..223642
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-201C19"
2773..3305
/notes="clone boundary
clone end:T7
site:EcoRI
end_sequence:BH330265"
17785..18888
/notes="wgs_contig"

misc_feature
2773..3305
/notes="clone boundary
clone end:T7
site:EcoRI
end_sequence:BH330265"

misc_feature
17785..18888
/notes="wgs_contig"

```

misc\_feature 135936 ..138590  
/note="wgs\_contig"

ORIGIN

Query Match 63.6%; Score 21; DB 2; Length 223642;  
Best Local Similarity 82.8%; Pred. No. 1.6e+02;  
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CCATTATACCTAGTTCCTGGTTGCT 30  
||||| | ||||| | ||||| |  
DB 137906 CCATTAGTTGTAGTACTGGTTGCT 137934

RESULT 37  
AC122632/c  
LOCUS  
DEFINITION Rattus norvegicus clone CH230-136C2, \*\*\* SEQUENCING IN PROGRESS  
ACCESSION AC122632 GI:30579821  
VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 236824)  
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Ayodeji,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhatz,C., Burck,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K., Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gregorogis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,K., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovac,C., Kowals,C., Kraft,C., Kraft,C., Lebow,H., Levay,J., Lewis,L., Li.Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuheva,L., Loulseege,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangun,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNell,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okwuonu,G., Olarpunsaagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,B., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shadman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smajls,D., Steidle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willison,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,



Mon Jun 6 15:23:31 2005

Db 227344 CCATTAGTTGTTAGCTCTCGTTGCT 227372

```

RESULT 39
AC099662 246933 bp DNA linear HTG 09-OCT-2002
LOCUS Rattus norvegicus clone CH230-1D24, *** SEQUENCING IN PROGRESS ***
DEFINITION 4 unordered pieces.
ACCESSION AC099662 GI:23195793
VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 246933)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oman, F.R., Allen, C.,
Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaria, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Delavia, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Eamhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Foster, P., Frantz, P.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Garner, T., Garza, N., Gill, R.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsone, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
Oragunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, J.L., Quiles, M., Ren, Y.,
Rivas, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Tang, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Thomas, S.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vazquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczky, R., Woodson, S., Worley, K.,
Wu, C., Wu, Y., Xu, Y., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 246933)
Worley, K.C.
Direct Submission
Submitted (17-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 246933)
Worley, K.C.
Direct Submission
Submitted (09-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 19, 2002 this sequence version replaced gi:21553924.

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
sequence may extend beyond the ends of the clone and there may be
contigs that consist entirely of whole genome shotgun sequence only
reads. Both end sequences and whole genome shotgun sequence only
contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: TUPM
Center clone name: CH230-1D24
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 199368 bases at least Q40
Consensus quality: 204132 bases at least Q30
Consensus quality: 206995 bases at least Q20
Estimated insert size: 233699; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 239159: contig of 239159 bp in length
* 239160 239259: gap of unknown length
* 239260 242864: contig of 3605 bp in length
* 242865 242964: gap of unknown length
* 242965 243976: contig of 1012 bp in length
* 243977 244076: gap of unknown length
* 244077 246933: contig of 2857 bp in length.
-----
FEATURES
Location/Qualifiers
1. 246933
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-1D24"
1..1303
misc_feature /note="wgs contig"
6946..7963
misc_feature /note="wgs contig"
239260..240869
misc_feature /note="wgs contig"

ORIGIN
Query Match 63.6%; Score 21; DB 2; Length 246933;
Best Local Similarity 82.8%; Pred. No. 1.6e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CCATTTATTACGCTAGTTCCTGTTGCT 30
|||||
Db 230533 CCATTAGTTGTTAGCTCTCGTTGCT 230561
|||||

RESULT 40
AC130059/c
LOCUS Rattus norvegicus clone CH230-14F14, *** SEQUENCING IN PROGRESS
DEFINITION AC130059
ACCESSION AC130059
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
1 (bases 1 to 246933)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oman, F.R., Allen, C.,
Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaria, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Delavia, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Eamhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Foster, P., Frantz, P.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Garner, T., Garza, N., Gill, R.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsone, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
Oragunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, J.L., Quiles, M., Ren, Y.,
Rivas, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Tang, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Thomas, S.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vazquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczky, R., Woodson, S., Worley, K.,
Wu, C., Wu, Y., Xu, Y., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 246933)
Worley, K.C.
Direct Submission
Submitted (17-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 246933)
Worley, K.C.
Direct Submission
Submitted (09-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 19, 2002 this sequence version replaced gi:21553924.

```

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GXND  
Center clone name: CH230-14F14  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 183816 bases at least Q40  
Consensus quality: 187963 bases at least Q30  
Estimated insert size: 190821 bases at least Q20  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 98797: contig of 98797 bp in length  
98798 98897: gap of unknown length  
98898 116591: contig of 17694 bp in length  
116592 116691: gap of unknown length  
116692 132445: contig of 15754 bp in length  
132446 132545: gap of unknown length  
132546 172419: contig of 39874 bp in length  
172420 172519: gap of unknown length  
172520 234512: contig of 61993 bp in length  
234513 234612: gap of unknown length  
234613 244555: contig of 9943 bp in length  
244556 245754: contig of unknown length  
245755 245855: gap of unknown length  
245856 245975: contig of 1121 bp in length.

1. 245975  
Location/Qualifiers  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-14F14"  
/note="wgs\_contig"  
97229. .98797

misc\_feature  
misc\_feature  
ORIGIN

Query Match 63.6%; Score 21; DB 2; Length 246975;  
Best Local Similarity 82.8%; Pred. No. 1.6e+02;  
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 2 CCATTATTACGTAGTTCCTGCTTTCCT 30  
Db 80640 CCATTATTAGTTGTTAGTCTGCTTTCCT 80612

Search completed: June 4, 2005, 14:31:42  
Job time : 753.667 secs

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Murida.; Murinae;  
1 (bases 1 to 246975)  
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Bacon, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, W., Benahmed, F.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, J., Carter, K., Cavazos, I., Ceasar, H., Chen, R.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
David, M., Davis, S., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, C. A., Foster, M., Foster, P.,  
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hollins, B., Howells, S., Hladun, S. L., Hodgson, A., Hojues, M.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, E., Jolivet, A.,  
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorenshewa, L., Loulseghe, H., Lozano, R. J., Lu, X., Me, J.,  
Mangum, B., Mapua, P., Martin, K., Mahmoud, M., Malloy, K., Martin, R.,  
Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwaokemele, O., Okwunonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,  
Rizzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,  
Rives, C., Rodkey, T., Rojce, A., Rose, M., Rose, R., Ruiz, S. J.,  
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,  
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,  
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,  
Weinstock, G. and Gibbs, R. A.  
Direct Submission  
Unpublished  
2 (bases 1 to 246975)  
Worley, K. C.  
Direct Submission  
Submitted (08-AUG-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 246975)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Nov 15, 2002 this sequence version replaced gi:2332277.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,

REFERENCE  
AUTHORS

## COMMENT



us-10-089-452-26.rng

Mon Jun 6 15:23:31 2005

Abal8630 Human ner  
Abd32649 Human onc  
Abz59002 Human onc  
Abz59000 Human onc  
Abz59003 Human onc  
Abz59008 Human onc  
Abz58998 Human onc  
Abz58996 Human onc  
Abz58995 Human onc  
Adq97643 Human onc  
Aag51631 Anti-lyso  
Abn70767 Streptoco  
Abn66547 Streptoco  
Aav74732 Staphyloc  
Adr84541 Aspergill  
Aad47150 Human Ras  
Aac44168 Mouse gen  
Abn71527 Streptoco  
Aac43998 Human gen  
Adq18503 Human foe  
Aah94452 Human foe  
Aad48017 Bacterial  
Abi05662 Drosophil  
Continuation (5 of  
Continuation (4 of  
Continuation (2 of

21 20.2 61.2 8668 5 ABal8630  
22 20.2 61.2 102790 13 ABD32649  
23 20 60.6 1203 8 ABZ59002  
24 20 60.6 1203 8 ABZ59000  
25 20 60.6 1637 8 ABZ59003  
26 20 60.6 2464 8 ABZ58998  
27 20 60.6 3761 8 ABZ58996  
28 20 60.6 10735 8 ABZ58995  
29 20 60.6 100137 12 ADQ97643  
30 19.8 60.0 369 2 AAG51631  
31 19.8 60.0 966 6 ABN70767  
32 19.8 60.0 978 6 ABN66547  
33 19.8 60.0 2710 2 AAV74732  
34 19.8 60.0 7179 13 ADR84541  
35 19.8 60.0 31814 10 AAD47150  
36 19.8 60.0 39121 11 ACN44168  
37 19.8 60.0 110000 6 ABN71527\_00  
38 19.8 60.0 110000 11 ACN43998\_0  
39 19.6 59.4 310 12 ADQ18503  
40 19.4 58.8 617 5 AAH94452  
41 19.4 58.8 903 13 ADS48017  
42 19.4 58.8 8361 4 ABL05662  
43 19.4 58.8 110000 6 ABA90521\_04  
44 19.4 58.8 110000 13 ABD32629\_3  
45 19.4 58.8 110000 13 ABD32806\_1

ALIGNMENTS

RESULT 1  
AAF88110  
ID AAF88110 standard; DNA; 33 BP.  
XX  
AC AAF88110;  
XX  
DT 17-JUL-2001 (first entry)  
XX  
DE H. pylori catalase derived antibody HP25/6m/1B5 heavy chain CDR3 DNA.  
XX  
KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;  
KW acid-resistant microorganism; complementarity determining region; CDR;  
KW feces; heavy chain; light chain; ds.  
XX  
OS Unidentified.  
XX  
PN WO200127612-A2.  
XX  
PD 19-APR-2001.  
XX  
PF 12-OCT-2000; 2000WO-EP010057.  
XX  
PR 12-OCT-1999; 99EP-00120351.  
PR 16-MAR-2000; 2000EP-00105592.  
PR 31-MAR-2000; 2000EP-00107028.  
PR 10-MAY-2000; 2000EP-00110110.  
(CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;  
WPI; 2001-282086/29.  
P-PSDB; AAB86083.  
Detecting infections by acid-resistant microorganisms, particularly for  
diagnosing Helicobacter pylori, comprises immunochromatographic detection  
of antigen in feces.  
Claim 26; Page 25; 90pp; German.  
This invention describes a novel method for detecting infection by an  
acid-resistant microorganism (A), in a mammal, using  
immunochromatography. The method is used to diagnose infection by an acid

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: June 4, 2005, 10:23:32 ; Search time 203.333 Seconds  
(without alignments)  
960.745 Million cell updates/sec

Title: US-10-089-452-26  
Perfect score: 33  
Sequence: 1 cccattattacgtagtctcgttggttcttac 33

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:.\*  
1: Geneseq1980s:.\*  
2: Geneseq1990s:.\*  
3: Geneseq2000s:.\*  
4: Geneseq2001as:.\*  
5: Geneseq2001bs:.\*  
6: Geneseq2002as:.\*  
7: Geneseq2002bs:.\*  
8: Geneseq2003as:.\*  
9: Geneseq2003bs:.\*  
10: Geneseq2003cs:.\*  
11: Geneseq2003ds:.\*  
12: Geneseq2004as:.\*  
13: Geneseq2004bs:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	33	4 AAF88110	Aaf88110 H. pylori
2	33	100.0	33	4 AAF88053	Aaf88053 H. pylori
3	33	100.0	360	4 AAF88153	Aaf88153 H. pylori
4	33	100.0	360	4 AAF88096	Aaf88096 H. pylori
5	23	69.7	345	2 AAV35191	Aav35191 Human H c
6	21.4	64.8	469	2 AAO12060	Aao12060 Sequence
7	21.4	64.8	469	2 AAO12016	Aao12016 Sequence
8	21.4	64.8	161671	12 ADH30232	Adh30232 Human PLO
9	20.8	63.0	3501	3 AAC79616	Aac79616 Virulence
10	20.8	63.0	3501	6 ABQ83493	Abq83493 Pasteurel
11	20.4	61.8	100	8 ACD77544	Acd77544 E. coli K
12	20.4	61.8	100	8 ACD77544	Acd77544 E. coli K
13	20.4	61.8	617	8 ACAL4876	Acac4876 Prokaryot
14	20.4	61.8	739	2 AAX03015	Aax03015 Human IL-
15	20.4	61.8	772	4 AAL25021	Aal25021 Human bre
16	20.4	61.8	1240	12 ADI42223	Adi42223 Plant tra
17	20.4	61.8	1240	12 ADO02726	Ado02726 Corn orth
18	20.4	61.8	1488	8 ACAL19191	Acac19191 Prokaryot
19	20.4	61.8	5362	2 AAX02980	Aax02980 Human IL-
20	20.2	61.2	1006	10 ADF41968	Adf41968 Alcaligen

CC -resistant microorganism (A), in a mammal, such as Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,  
 CC inexpensive and non-invasive, and may indicate the stage of infection. A  
 CC test strip used in the method may include a filter to eliminate particles  
 CC present in the sample and only a single receptor provides a reasonably  
 CC secure diagnosis, with specificity and selectivity improved by detecting  
 CC several epitopes (of catalase) or different antigens (catalase and beta-  
 CC urease). The method can be automated. This sequence encodes a  
 CC complementarity determining region (CDR) from an antibody raised against  
 CC the H. pylori catalase or beta-urease antigen which is used to illustrate  
 CC the method of the invention  
 CC  
 XX Sequence 33 BP; 5 A; 9 C; 5 G; 14 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 33; DB 4; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 0.00034;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CCCATTATTACGCTAGTCTCTGGTTGCTTAC 33  
 Db 1 CCCATTATTACGCTAGTCTCTGGTTGCTTAC 33  
 RESULT 2  
 AAF88053  
 ID AAF88053 standard; DNA; 33 BP.  
 AC AAF88053;  
 XX  
 DT 17-JUL-2001 (first entry)  
 DE H. pylori catalase derived antibody HP25/6m/1B5 heavy chain CDR3 DNA.  
 KW Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;  
 KW infection; acid-resistant microorganism; fecal; antibody; diagnosis;  
 KW antibacterial; complementarity determining region; ds.  
 XX Unidentified.  
 OS  
 XX WO200127613-A2.  
 PN  
 XX 19-APR-2001.  
 PD  
 XX 12-OCT-2000; 2000WO-EP010058.  
 PF  
 XX 12-OCT-1999; 99EP-00120351.  
 PR 16-MAR-2000; 2000EP-00105592.  
 PR 31-MAR-2000; 2000EP-00107028.  
 PR 10-MAY-2000; 2000EP-00110110.  
 XX  
 PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
 XX  
 PI Reiter C, Cullmann G, Heppner P, Ringels A, Mueller H, Haindl E;  
 XX  
 DR WPI; 2001-282087/29.  
 DR P-PSDB; AAB86051.  
 XX  
 PT Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing Helicobacter pylori, comprises an immunoassay on a fecal  
 PT sample.  
 XX  
 PS Claim 20; Page 16; 89pp; German.

CC synthetic peptides). The products of the invention have antibacterial  
 CC activity. The method is used to diagnose infection by Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the  
 CC progress of treatment. Receptors; particularly antibodies, directed  
 CC against Ag can be used therapeutically for treatment of infections. The  
 CC method requires only one R to provide a reasonably secure diagnosis.  
 CC (although use of two R improves sensitivity), so is relatively  
 CC inexpensive and more easily standardized. Also it is direct, non-  
 CC invasive, suitable for automation and may indicate the stage of an  
 CC infection. This sequence encodes a complementarity determining region  
 CC (CDR) from an antibody generated against a Helicobacter pylori antigen  
 CC (catalase or beta-urease) which is used to illustrate the method of the  
 CC invention  
 CC  
 XX Sequence 33 BP; 5 A; 9 C; 5 G; 14 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 33; DB 4; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 0.00034;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CCCATTATTACGCTAGTCTCTGGTTGCTTAC 33  
 Db 1 CCCATTATTACGCTAGTCTCTGGTTGCTTAC 33  
 RESULT 3  
 AAF88153  
 ID AAF88153 standard; DNA; 360 BP.  
 AC AAF88153;  
 XX  
 DT 17-JUL-2001 (first entry)  
 DE H. pylori catalase antibody HP25/6m/1B50 V region heavy chain DNA.  
 KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;  
 KW acid-resistant microorganism; complementarity determining region; CDR;  
 KW feces; heavy chain; light chain; ds.  
 XX Unidentified.  
 OS  
 XX WO200127612-A2.  
 PN  
 XX 19-APR-2001.  
 PD  
 XX 12-OCT-2000; 2000WO-EP010057.  
 PF  
 XX 12-OCT-1999; 99EP-00120351.  
 PR 16-MAR-2000; 2000EP-00105592.  
 PR 31-MAR-2000; 2000EP-00107028.  
 PR 10-MAY-2000; 2000EP-00110110.  
 XX  
 PA (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.  
 XX  
 PI Reiter C, Cullmann G, Lakner M, Truee A, Dehnert S, Schwartz G;  
 XX  
 DR WPI; 2001-282086/29.  
 DR P-PSDB; AAB86107.  
 XX  
 PT Detecting infections by acid-resistant microorganisms, particularly for  
 PT diagnosing Helicobacter pylori, comprises immunochromatographic detection  
 PT of antigen in feces.  
 XX  
 PS Claim 32; Fig 3; 90pp; German.

CC This invention describes a novel method for detecting infection by an  
 CC acid-resistant microorganism (A), in a mammal, using  
 CC immunochromatography. The method is used to diagnose infection by an  
 CC -resistant microorganism (A), in a mammal, such as Helicobacter,  
 CC Campylobacter or Mycobacterium, particularly H. pylori (most preferred),  
 CC H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,  
 CC inexpensive and non-invasive, and may indicate the stage of infection A



QY 3 CATTATTACGCTAGTTCCTGCTTGTCTTAC 33  
 |||||  
 Db 279 CATTATTACTGCTAGTTCCTGCTTGTCTTAC 309  
 |||||

## RESULT 6

AAQ12060  
 ID AAQ12060 standard; DNA; 469 BP.  
 XX AC AAQ12060;  
 XX 25-MAR-2003 (revised)  
 DT 15-AUG-1991 (first entry)  
 XX  
 XX Sequence encoding heavy chain variable region of murine 1C11  
 DE immunoglobulin.  
 DE  
 XX Chimeric antibodies; immunoconjugates; HIV; AIDS; ss.  
 KW  
 XX Mus musculus.

XX Key Location/Qualifiers  
 FH CDS 67..468  
 FT /\*tag= a  
 FT /product= "mouse MAB 1C11 H-chain variable region"  
 XX  
 XX WO9107493-A.

XX 30-MAY-1991.  
 XX 13-NOV-1989; 89US-00433730.  
 XX 13-NOV-1989; 89US-00433730.  
 XX (XOMA ) XOMA CORP.  
 PA (GREG ) GREEN CROSS CORP.

XX Better MD, Horwitz AH, Ghoshdasti P, Robinson R;  
 WPI; 1991-178105/24.  
 DR P-PSDB; AAR12356.  
 XX  
 XX New chimeric mouse-human antibodies - used to detect, kill and remove HIV  
 PT -1 antigen from sample.  
 XX  
 XX Disclosure; Fig 12; 107pp; English.

XX This sequence encodes the heavy chain variable (V) region of a mouse  
 CC monoclonal antibody (MAB), 1C11, specific for an HIV-1 viral antigen. It  
 CC is used in the construction of a chimeric MAB comprising heavy and light  
 CC chains having murine V regions and human C regions. The chimeric MABs are  
 CC more effective than murine MAB 1C11 since they have an increased  
 CC compatibility in humans. The heavy and light chain V-regions are joined  
 CC by manipulating their respective joining (J) regions, to generate  
 CC restriction enzyme recognition sites. The chimeric MABs can be used as  
 CC immuno- conjugates, in association with e.g. toxins for HIV treatment..  
 CC They can also be used in diagnosis of HIV. See also AAQ12056-59 and  
 CC AAQ12061-63. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25  
 CC -MAR-2003 to correct PI field.) (Updated on 25-MAR-2003 to correct DR  
 CC field.)  
 XX

SQ Sequence 469 BP; 127 A; 96 C; 120 G; 126 T; 0 U; 0 Other;

Query Match 64.8%; Score 21.4; DB 2; Length 469;  
 Best Local Similarity 80.6%; Pred. No. 38;  
 Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGCTTGTCTTAC 33  
 |||||  
 Db 405 CATTATTACTGCTAGTTCCTGCTTGTCTTAC 435  
 |||||

## RESULT 7

AAQ12016  
 ID AAQ12016 standard; DNA; 469 BP.  
 XX AC AAQ12016;  
 XX 25-MAR-2003 (revised)  
 DT 19-AUG-1991 (first entry)  
 XX  
 XX Sequence encoding mouse MAB 1C11 H chain V region.

DE HIV-1; chimera; ds.

XX Mus sp.

XX Key Location/Qualifiers  
 FH CDS 67..469  
 FT /\*tag= a

XX WO9107494-A.

XX 30-MAY-1991.

XX 13-NOV-1989; 89US-00433703.

XX 13-NOV-1989; 89US-00433703.

XX (XOMA ) XOMA CORP.

PA (GREG ) GREEN CROSS CORP.

PA (ZOMA-) ZOMA CORP.

XX Better MD, Horwitz AH, Ghoshdasti P, Robinson RR;

XX WPI; 1991-178106/24.

DR P-PSDB; AAR12236.

XX New chimeric mouse human antibodies - used in treatment, diagnosis and  
 PT prophylaxis of HIV infections.

XX Disclosure; Fig 12; 108pp; English.

XX The mouse VH gene product may be used to produce chimeric mouse- human  
 CC Abs against HIV-1 comprising human Ig constant regions and murine  
 CC variable regions. These novel sequence are useful in treatment, diagnosis  
 CC and prophylaxis of HIV infections, and may be produced by a bacterial,  
 CC yeast or mammalian expression system. (Updated on 25-MAR-2003 to correct  
 CC PA field.) (Updated on 25-MAR-2003 to correct PI field.)  
 XX

SQ Sequence 469 BP; 127 A; 96 C; 120 G; 126 T; 0 U; 0 Other;

Query Match 64.8%; Score 21.4; DB 2; Length 469;  
 Best Local Similarity 80.6%; Pred. No. 38;  
 Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGCTTGTCTTAC 33  
 |||||  
 Db 405 CATTATTACTGCTAGTTCCTGCTTGTCTTAC 435  
 |||||

## RESULT 8

ADH30232  
 ID ADH30232 standard; DNA; 161671 BP.

XX ADH30232;

XX 11-MAR-2004 (first entry)

XX Human PLOD2 DNA.

XX SNP; single nucleotide polymorphism; ds; gene; vascular disease; PLOD2;  
 KW atherosclerosis; coronary artery disease; myocardial infarction;  
 KW ischaemia; stroke; peripheral vascular disease; venous thromboembolism;  
 KW pulmonary embolism; paternity testing; human.  
 XX

```
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT replace(147472.g)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
XX
XX US2003124535-A1.
XX
XX 03-JUL-2003.
XX
XX 14-DEC-2001; 2001US-00017117.
XX
XX 05-DEC-2001; 2001US-0337883P.
XX
XX (VITI-) VITIVITY INC.
XX
XX Mccarthy J;
XX
XX WPI; 2004-068891/07.
XX
XX P-PSDB; ADH30233.
XX
XX Methods for determining the PLOD2 genetic profile of a subject based on
XX the discovery of a polymorphism within the lysyl hydroxylase 2 (PLOD2)
XX gene for diagnosing or aiding in the diagnosis of a vascular disease or a
XX disorder in a subject.
XX
XX Claim 1; SEQ ID NO 1; 155pp; English.
XX
XX The invention relates to a method for diagnosing or aiding in the
XX diagnosis of a vascular disease or a disorder in a subject comprising
XX determining the PLOD2 genetic profile of the subject. Methods are
XX provided for diagnosing or aiding in the diagnosis of a vascular disease
XX or a disorder in a subject e.g. atherosclerosis, coronary artery disease,
XX myocardial infarction, ischaemia, stroke, peripheral vascular diseases,
XX venous thromboembolism and pulmonary embolism. The invention also
XX provides methods for classifying a subject who is at risk or not for
XX developing a vascular disease or disorder as a candidate for a particular
XX course of therapy or a particular diagnostic evaluation. The invention
XX further provides methods for selecting a clinical course of therapy or a
XX diagnostic evaluation to treat a subject who is at risk or not for
XX developing a vascular disease or disorder. The nucleic acids of the
XX invention can be used in forensics for determining polymorphisms and in
XX paternity testing. The invention provides an early means of detecting
XX vascular disease and ways of preventing the full scale development of
XX vascular diseases or disorders. The present sequence represents human
XX PLOD2 DNA.
XX
XX SQ Sequence 161671 BP; 50834 A; 28898 C; 28646 G; 52188 T; 0 U; 1105 Other;
XX
XX Query Match 64.8%; Score 21.4; DB 12; Length 161671;
XX Best Local Similarity 80.6%; Pred. No. 80;
XX Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 1 CCCATTATTACGCTAGTTCCTGGTTGCTT 31
XX ||||| ||||| ||||| ||||| |||||
XX 125929 CCCATTAAACTCTATTTCATTGTTGCTT 125959
XX
XX RESULT 9
XX AAC79616
XX ID AAC79616 standard; DNA; 3501 BP.
XX
XX AC AAC79616;
XX
XX 08-FEB-2001 (first entry)
XX
XX Virulence gene #36.
XX
XX Virulence gene; antibacterial; vaccine; bacterial infection; septicemia;
XX bronchopneumonia; rhinitis; wound infection; ss.
XX
XX Pasteurella multocida.
XX
```

```
XX WO200061724-A2.
XX
XX 19-OCT-2000.
XX
XX 06-APR-2000; 2000WO-US009218.
XX
XX 09-APR-1999; 99US-0128689P.
XX
XX 10-SEP-1999; 99US-0153453P.
XX
XX (PHAA ) PHARMACIA & UPJOHN INC.
XX
XX Lowery DE, Fuller TE, Kennedy MJ;
XX
XX WPI; 2000-647422/62.
XX
XX P-PSDB; AAB44555.
XX
XX Attenuated Pasteurellaceae bacteria comprising mutations in virulence
XX genes, useful as a live attenuated vaccine against bacterial infections.
XX
XX Claim 7; Page 204-207; 322pp; English.
XX
XX The family Pasteurellaceae encompasses several pathogens that infect a
XX wide variety of animals. The present invention relates to virulence genes
XX from Pasteurellaceae. The present sequence is one such virulence gene.
XX The present sequence may be mutated in order to produce an inactive gene.
XX The inactive virulence gene may in turn be used to produce a vaccine,
XX which is useful for treating bacterial infections such as septicemias,
XX bronchopneumonias, rhinitis and wound infections
XX
XX SQ Sequence 3501 BP; 896 A; 684 C; 801 G; 1120 T; 0 U; 0 Other;
XX
XX Query Match 63.0%; Score 20.8; DB 3; Length 3501;
XX Best Local Similarity 78.1%; Pred. No. 87;
XX Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX QY 2 CCATTATTACGCTAGTTCCTGGTTGCTTAC 33
XX ||||| ||||| ||||| ||||| |||||
XX 1216 CTACTCAATACGCTAGTTCCTGGTTGCTTTC 1247
XX
XX RESULT 10
XX ABQ83493
XX ID ABQ83493 standard; DNA; 3501 BP.
XX
XX AC ABQ83493;
XX
XX 07-AUG-2003 (revised)
XX
XX 24-JAN-2003 (first entry)
XX
XX Pasteurella multocida yabk gene SEQ ID NO:70.
XX
XX Antibacterial; vaccine; gram negative bacterial virulence gene;
XX identification; virulence; Pasteurellaceae; gene; ds.
XX
XX Pasteurella multocida.
XX
XX WO200275507-A2.
XX
XX 26-SEP-2002.
XX
XX 17-JAN-2002; 2002WO-US001971.
XX
XX 15-MAR-2001; 2001US-00809665.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Lowery DE, Fuller TE, Kennedy MJ;
XX
XX WPI; 2002-740868/80.
XX
XX P-PSDB; ABP54507.
XX
XX New mutant gram-negative bacteria, useful as vaccines and for identifying
XX
```

PT new anti-bacterial agents that target virulence genes and their products.  
 PS Claim 5; Page 209-212; 350pp; English.

XX The present invention describes a gram-negative bacteria comprising a  
 CC mutation in a gene, where the mutation results in decreased activity of a  
 CC gene product encoded by the mutated gene. Also described is a method for  
 CC producing a gram-negative bacteria mutant or an attenuated  
 CC Pasteurellaceae bacteria. The mutated genes have antibacterial activity  
 CC and can be used in vaccines. The gram-negative bacteria or the attenuated  
 CC Pasteurellaceae bacteria can be used as vaccines in the fields of human  
 CC medicine or veterinary medicine, and for identifying new antibacterial  
 CC agents that target the virulence genes and their products. ABQ83458 to  
 CC ABQ83578 and ABP54473 to ABP54551 represents sequences used in the  
 CC exemplification of the present invention. (Updated on 07-AUG-2003 to  
 CC correct OS field.)

XX SQ Sequence 3501 BP; 896 A; 684 C; 801 G; 1120 T; 0 U; 0 Other;  
 Query Match 63.0%; Score 20.8; DB 6; Length 3501;  
 Best Local Similarity 78.1%; Pred. No. 87;  
 Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2 CCAATTATTACGCTAGTTCTCTGTTGCTTAC 33  
 Db 1216 CTACTCATACGCTAGTTCTCTGTTGCTTTC 1247

RESULT 11  
 ACD77543/c  
 ID ACD77543 standard; DNA; 100 BP.  
 XX AC ACD77543;  
 DT 19-SEP-2003 (first entry)  
 XX E. coli K12 MG1655 biochip probe SEQ ID 8819.  
 DE Biochip; gene expression; gut; diagnostic; detection; probe; ss.  
 KW Escherichia coli.  
 XX EP1260592-A1.  
 XX 27-NOV-2002.  
 XX 17-MAY-2001; 2001EP-00112179.  
 XX 17-MAY-2001; 2001EP-00112179.  
 XX (MWGB-) MWG-BIOTECH AG.  
 XX Donner H, Drescher B, Huber A, Weber J;  
 XX WPI; 2003-241155/24.

PS Biochip containing probes complementary with open reading frames in  
 PT Escherichia coli K12, useful for detecting gene expression and expression  
 PT patterns.  
 XX Claim 3; Page 1376; 2004pp; German.

XX This invention describes a novel biochip comprising probe spots, each  
 CC containing many identical probes. The probes are nucleotide sequences of  
 CC 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at  
 CC least one includes a segment of at least 20 bases identical with, or  
 CC complementary to, a segment of an open reading frame (orf) of Escherichia  
 CC coli K12. The biochip is used for specific detection of gene expression  
 CC in K12 and for determining the gene expression pattern, e.g. for  
 CC diagnostic determination of which E. coli strains are present in the gut,  
 CC and to determine the effects of e.g. growth media on gene expression. The  
 CC biochip provides a comprehensive as possible detection of the K12  
 CC genome, with simultaneous analysis of many different genes with a single

CC device, and comparison of gene expression between K12 and its mutants or  
 CC other E. coli strains in a single experiment. Apart from qualitative and  
 CC quantitative information about gene expression, it also allows the use of  
 CC measurements of population densities for the various strains. The use of  
 CC synthetic oligonucleotides for preparation of probes allows free  
 CC variation in probe length and ensures high purity (and thus selectivity,  
 CC reactivity and reproducibility); also synthetic probes are generally  
 CC shorter than probes prepared by polymerase chain reaction. ACD68731 to  
 CC ACD81540 represent oligonucleotide probes used with the biochip described  
 CC in the invention

XX SQ Sequence 100 BP; 30 A; 26 C; 23 G; 21 T; 0 U; 0 Other;  
 Query Match 61.8%; Score 20.4; DB 8; Length 100;  
 Best Local Similarity 80.0%; Pred. No. 82;  
 Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 ATTTATTACGCTAGTTCTCTGTTGCTTAC 33  
 Db 59 AGTCGTTACGCTCGTTACAGGTTGCTTAC 30

RESULT 12  
 ACD77544/c  
 ID ACD77544 standard; DNA; 100 BP.

XX AC ACD77544;  
 XX 19-SEP-2003 (first entry)  
 XX E. coli K12 MG1655 biochip probe SEQ ID 8820.  
 DE Biochip; gene expression; gut; diagnostic; detection; probe; ss.  
 KW Escherichia coli.  
 XX EP1260592-A1.  
 XX 27-NOV-2002.  
 XX 17-MAY-2001; 2001EP-00112179.  
 XX 17-MAY-2001; 2001EP-00112179.  
 XX (MWGB-) MWG-BIOTECH AG.  
 XX Donner H, Drescher B, Huber A, Weber J;  
 XX WPI; 2003-241155/24.

PS Biochip containing probes complementary with open reading frames in  
 PT Escherichia coli K12, useful for detecting gene expression and expression  
 PT patterns.

XX Claim 3; Page 1376; 2004pp; German.

XX This invention describes a novel biochip comprising probe spots, each  
 CC containing many identical probes. The probes are nucleotide sequences of  
 CC 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at  
 CC least one includes a segment of at least 20 bases identical with, or  
 CC complementary to, a segment of an open reading frame (orf) of Escherichia  
 CC coli K12. The biochip is used for specific detection of gene expression  
 CC in K12 and for determining the gene expression pattern, e.g. for  
 CC diagnostic determination of which E. coli strains are present in the gut,  
 CC and to determine the effects of e.g. growth media on gene expression. The  
 CC biochip provides a comprehensive as possible detection of the K12  
 CC genome, with simultaneous analysis of many different genes with a single  
 CC device, and comparison of gene expression between K12 and its mutants or  
 CC other E. coli strains in a single experiment. Apart from qualitative and  
 CC quantitative information about gene expression, it also allows the use of  
 CC measurements of population densities for the various strains. The use of  
 CC synthetic oligonucleotides for preparation of probes allows free  
 CC variation in probe length and ensures high purity (and thus selectivity,

CC reactivity and reproducibility); also synthetic probes are generally  
CC shorter than probes prepared by polymerase chain reaction. ACD68731 to  
CC ACD61540 represent oligonucleotide probes used with the biochip described  
CC in the invention

XX Sequence 100 BP; 29 A; 24 C; 25 G; 22 T; 0 U; 0 Other;

Query Match 61.8%; Score 20.4; DB 8; Length 100;

Best Local Similarity 80.0%; Pred. No. 82; Mismatches 6; Indels 0; Gaps 0;

Matches 24; Conservative 0;

QY 4 ATTTATTACGCTAGTTCCTGGTTGCTTAC 33

Db 72 AGTCGTACGCTCGTTACAGGTTTGCTTAC 43

RESULT 13

ACA14876

ID ACA14876 standard; DNA; 617 BP.

XX ACA14876;

AC ACA14876;

XX 27-OCT-2003 (revised)

DT 19-JUN-2003 (first entry)

XX Prokaryotic essential gene antisense oligonucleotide #2746.

DE Antisense; ss; prokaryotic essential gene; cell proliferation;  
KW drug design.

XX Archaea.

OS Archaea.

XX WO200277183-A2.

PN 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

PI WPI; 2003-029926/02.

DR New antisense nucleic acids, useful for identifying proteins or screening

XX for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 1; SEQ ID NO 2746; 1766pp; English.

PS The invention relates to an isolated nucleic acid comprising any one of

XX the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the 6213  
CC antisense sequences of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 27-OCT-2003 to  
CC standardise OS field)

XX Sequence 617 BP; 132 A; 174 C; 149 G; 162 T; 0 U; 0 Other;

Query Match 61.8%; Score 20.4; DB 8; Length 617;

Best Local Similarity 80.0%; Pred. No. 16+02;

Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 ATTTATTACGCTAGTTCCTGGTTGCTTAC 33

Db 107 AGTCGTACGCTCGTTACAGGTTTGCTTAC 136

RESULT 14

AA03015/c

ID AA03015 standard; DNA; 739 BP.

XX AA03015;

AC AA03015;

XX 22-JUN-1999 (first entry)

XX Human IL-1ra BAC contiguous DNA sequence 60.

XX Tango-77; human; IL-1ra; cytokine superfamily; inflammation; inhibition;

XX interleukin-1 receptor; IL-1R; regulation; asthma; rheumatoid arthritis;

XX chronic myelogenous leukaemia; psoriasis; inflammatory bowel disease;

XX growth factors; treatment; IL-1 receptor complex; BAC; ss.

XX Homo sapiens.

OS WO9906426-A1.

XX 11-FEB-1999.

XX 03-AUG-1998; 98WO-US016102.

XX 04-AUG-1997; 97US-0054646P.

PR 02-JUL-1998; 98US-0091650P.

XX (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

PA Pan Y;

XX WPI; 1999-153692/13.

XX New isolated nucleic acid encoding the new human cytokine Tango-77 - used

XX to inhibit inflammation and to screen for specific modulators.

XX Example 5; Fig 4; 226pp; English.

XX AAX02956-X03048 and AAX22301-X22304 are overlapping BAC genomic sequences

XX containing alternatively spliced forms of human IL-1ra. Such fragments

XX are used in the method of the invention which describes the isolation of

XX a novel human TANGO-77 encoding nucleic acid and protein. Tango-77 is a

XX member of the cytokine superfamily that is expected to inhibit

XX inflammation by binding to the interleukin-1 receptor (IL-1R). It may

XX also bind to a new receptor so could regulate other cellular processes

XX associated with acute or chronic inflammation, e.g. asthma, chronic

CC myelogenous leukaemia, psoriasis and inflammatory  
CC bowel disease. It may also induce or suppress interleukins, cytokines and  
CC growth factors. Modulators of this protein are used to treat or prevent  
CC conditions associated with abnormal levels of inflammation, or activity  
CC of IL-1 or its receptor complex

SQ Sequence 739 BP; 235 A; 97 C; 136 G; 270 T; 0 U; 1 Other;  
 Query Match 61.8%; Score 20.4; DB 2; Length 739;  
 Best Local Similarity 80.0%; Pred. No. 1.1e+02;  
 Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 CCATTTATTACGCTAGTTCCTGGTTTGCTT 31  
Db 526 CCATCTATGAAGTTAAATTCCTGGTTTGCTT 497

RESULT 15  
AAL25021  
ID AAL25021 standard; cDNA; 772 BP.

DT	07-DEC-2001 (first entry)
XX	
DE	Human breast cancer expressed polynucleotide 17478.
XX	
KW	Human; breast cancer; cell marker; cytostatic; ss.

The invention relates to human breast cancer expressed polynucleotides (AA007544-AA026799) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterizing treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity

SQ Sequence 772 BP; 244 A; 122 C; 130 G; 276 T; 0 U; 0 Other;  
 Query Match 61.8%; Score 20.4; DB 4; Length 772;  
 Best Local Similarity 80.0%; Pred No. 1.1e+02;  
 Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 ATTTATTACGCTAGTTCCCTGGTTTGCTTAC 33

Db 540 ATTTTGTAGCTAGTACTGGTATCCTTAC 569

RESULT 16  
ADI42223/c  
ID ADI42223 standard; DNA; 1240 BP.

DT	22-APR-2004 (first entry)
XX	
DE	Plant transcription factor polynucleotide #403.

transgenic; plant; enhanced tolerance to abiotic stress;  
 glyphosate tolerance; hormone sensitivity; disease resistance;  
 sugar sensing; flowering; flower structure; stem bifurcation;  
 branching pattern; apical dominance; trichome; stem morphology;  
 root growth; root hair; seed development; cell proliferation;  
 cell differentiation; premature senescence; necrosis; plant size;  
 leaf morphology; seed morphology; seed biochemistry; root ancyonin;  
 plant ancyonin; light response; shade avoidance; bioinformatic;  
 transcription factor; gene; da.

OS Zea mays.

PN US2004019927-A1.

PD 29-JAN-2004.

25-FEB-2003; 2003US-00374780.

PR 18-APR-2001; 2001US-00837944.

PA (SHER/) SHERMAN B K.  
PA (RIEC/) RIECHMANN J L.

PA (HEAR//) HEARD J B

PA (CREE/) CREELMAN R A.

PA (ADAM/) ADAM L J.

PA (KEDD/) KEDDIE J.

PA (PILG/) PILGRIM M L

PA (PINE/) PINED,

PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;  
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;  
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;  
XX  
DR WPI; 2004-132345/13.

New transgenic plant comprising a recombinant polynucleotide of any one of more than 500 nucleotide sequences, useful in bioinformatic search methods.

PS Claim 1; SEQ ID NO 686; 435pp; English.

The invention describes a transgenic plant comprising a recombinant polynucleotide of any one of more than 500 nucleotide sequences fully defined in the specification or its complement. The method of the invention can be used to produce a plant having altered traits such as: enhanced tolerance to abiotic stress; glyphosate tolerance; hormone sensitivity; disease resistance; sugar sensing; early or late flowering; altered flower structure, change in stem bifurcations, altered branching pattern, reduced apical dominance, reduced trichome density; lack of trichomes; reduced ectopic trichome development; altered trichome development; increase in trichome number; altered stem morphology; increased root growth; increased root hairs; altered seed development; altered cell proliferation or cell differentiation; rapid development;



PA	(ADAM//) ADAM L J.
PA	(BROU//) BROUD P E.
PA	(PINE//) PINEDA O.
PA	(REUB//) REUBER T L.
PA	(KEDD//) KEDDIE J S.
PA	(YUGG//) YU G.
PA	(JIAN//) JIANG C.
PA	(SAMA//) SAMAH R S.
PA	(PILG//) PILGRIM M L.
PA	(CREE//) CREELMAN R A.
PA	(DUBE//) DUBELL A N.
PA	(RATC//) RATCLIFFE O.
PA	(KOMI//) KUMIMOTO R.
PA	(SHER//) SHERMAN B K.
XX	
XX	Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;
PI	Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;
PI	Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;
PI	Sherman BK;
XX	
XX	WPI; 2004-225755/21.
DR	
XX	
XX	New transgenic plant, useful in developing phenotypes with altered or
PT	improved characteristics or traits.
XX	
XX	Claim 1; SEQ ID NO 1140; 213pp; English.
XX	
CC	The invention relates to a transgenic plant comprises a recombinant
CC	polynucleotide having a polynucleotide sequence or its complementary
CC	sequence comprising a sequence encoding a polypeptide, that initiates
CC	transcription (i.e. a transcription factor) from Arabidopsis, Soybean,
CC	Rice, Rape or Corn, comprising any of the sequences appearing as ADO01588
CC	-ADO03537 or ADO03530-ADO03559. Also included are using a transgenic
CC	plant to grow a progeny plant, an expression cassette (comprising a
CC	constitutive, inducible or tissue-specific promoter and a recombinant
CC	polynucleotide described above), a host cell comprising the expression
CC	cassette, producing a modified plant having a modified trait, identifying
CC	a factor that is modulated by or interacts with a polypeptide encoded by
CC	the polynucleotide sequence and identifying at least one downstream
CC	polynucleotide sequence that is subject to a regulatory effect of any of
CC	the polypeptides encoded by the polynucleotide described above. The
CC	transgenic plant is useful for producing a plant that has an altered
CC	trait e.g. an enhanced tolerance to abiotic stress (increased tolerance
CC	to chilling, germination in cold conditions,freezing tolerance, tolerance
CC	to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
CC	salt, tolerance to phosphate limitation, tolerance to potassium
CC	limitation, decreased sensitivity to nitrogen limitation), altered
CC	hormone sensitivity, reduced sensitivity to abscisic acid, an altered
CC	response to ethylene, disease resistance, altered susceptibility to
CC	Botrytis, altered susceptibility to Fusarium, altered susceptibility to
CC	Erysiphe, altered susceptibility to Pseudomonas syringae, altered
CC	susceptibility to Sclerotinia, altered sugar sensing, improved seed
CC	germination and seedling vigor, early flowering, late flowering, extended
CC	period of flowering, an inflorescence architectural change, a change in
CC	stem bifurcations, a lack of a shoot meristem, reduced meristem cell
CC	differentiation, altered phyllotaxy, altered branching pattern, reduced
CC	apical dominance, reduced trichome density, ectopic trichome development,
CC	altered trichome development, altered stem morphology, increased root
CC	growth, increased root hairs, altered seed development, altered cell
CC	proliferation/cell differentiation, premature senescence, delayed
CC	senescence, lethality, increased necrosis, an increase in seedling or
CC	plant size, decreased plant size, a change in leaf morphology, increased
CC	altered leaf development, increased leaf size and mass, glossy leaves,
CC	leaf cell expansion, change in seed morphology, altered seed coloration,
CC	increased seed size, decreased seed size, altered seed shape, change in
CC	leaf biochemistry, increased leaf wax, an alteration in leaf prenly lipid
CC	content, increased leaf insoluble sugars, decreased leaf insoluble
CC	sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
CC	content, an alteration of leaf glucosinolate content, change in seed
CC	biochemistry, an increase in seed oil content, decrease in seed oil
CC	content, increase in seed fatty acid content, decrease in seed fatty acid
CC	content, increase in seed protein content, decrease in seed protein
CC	content, alteration in seed prenlylipid content, increase n seed

CC sterols, upregulation of genes involved in secondary metabolism, increase  
CC in root anthocyanins, increase in plant anthocyanins, and alteration in  
CC light response or shade avoidance. The present sequence encodes an  
CC orthologue of a thalass transcription factor isolated from Corn.

XX SQ Sequence 1240 BP; 280 A; 348 C; 372 G; 239 T; 0 U; 1 Other;

Query Match 61.8%; Score 20.4; DB 12; Length 1240;  
Best Local Similarity 80.0%; Pred. No. 1.1e+02;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCCATTATACGCTAGTCTCGGTTGCT 30  
DB 1101 CCCAGTTTAACTCTAGTCTCTAGTTGCT 1072

RESULT 18  
ACA19191/c  
ID ACA19191 standard; DNA; 1488 BP.

XX AC ACA19191;

DT 19-JUN-2003 (first entry)

XX DE Prokaryotic essential gene #848.

XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
XX KW drug design; gene.

XX OS Escherichia coli.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABU15321.

XX New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 7061; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
XX the 6213 antisense sequences given in the specification where expression  
XX of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX (1) a vector comprising a promoter operably linked to the nucleic acid  
XX encoding a polypeptide whose expression is inhibited by the antisense  
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX polypeptide or its fragment whose expression is inhibited by the  
XX antisense nucleic acid; (4) an antibody capable of specifically binding  
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX proliferation or the activity of a gene in an operon required for  
XX proliferation; (7) identifying a compound that influences the activity of  
XX the gene product or that has an activity against a biological pathway  
XX required for proliferation, or that inhibits cellular proliferation; (8)  
XX identifying a gene required for cellular proliferation or the biological  
XX pathway in which a proliferation-required gene or its gene product lies  
XX or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 1488 BP; 364 A; 373 C; 426 G; 325 T; 0 U; 0 Other;

Query Match 61.8%; Score 20.4; DB 8; Length 1488;  
Best Local Similarity 80.0%; Pred. No. 1.2e+02;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 ATTATTACGCTAGTCTCTGTTGCTTAC 33

DB 1447 AGTCGTTACGCTCGTTACAGGTTGCTTAC 1418

RESULT 19

AAx02980/c

ID AAX02980 standard; DNA; 5362 BP.

XX AC AAX02980;

XX DT 22-JUN-1999 (first entry)

XX DE Human IL-1ra BAC contiguous DNA sequence 25.

XX KW Tango-77; human; IL-1ra; cytokine superfamily; inflammation; inhibition;  
XX KW interleukin-1 receptor; IL-1R; regulation; asthma; rheumatoid arthritis;  
XX KW chronic myelogenous leukaemia; psoriasis; inflammatory bowel disease;  
XX KW growth factors; treatment; IL-1 receptor complex; BAC; ss.

XX OS Homo sapiens.

XX PN WO9906426-A1.

XX PD 11-FEB-1999.

XX PF 03-AUG-1998; 98WO-US016102.

XX PR 04-AUG-1997; 97US-0054646P.

XX PR 02-JUL-1998; 98US-0091650P.

XX PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

XX PI Pan Y;

XX WPI; 1999-153692/13.

XX New isolated nucleic acid encoding the new human cytokine Tango-77 - used  
XX to inhibit inflammation and to screen for specific modulators.

XX Example 5; Fig 3; 226pp; English.

XX AAX02956-X03048 and AAX22301-X22304 are overlapping BAC genomic sequences  
XX containing alternatively spliced forms of human IL-1ra. Such fragments  
XX are used in the method of the invention which describes the isolation of  
XX a novel human TANGO-77 encoding nucleic acid and protein. Tango-77 is a  
XX member of the cytokine superfamily that is expected to inhibit  
XX inflammation by binding to the interleukin-1 receptor (IL-1R). It may  
XX also bind to a new receptor so could regulate other cellular processes  
XX associated with acute or chronic inflammation, e.g. asthma, chronic  
XX myelogenous leukaemia, rheumatoid arthritis, psoriasis and inflammatory

CC bowel disease. It may also induce or suppress interleukins, cytokines and  
CC growth factors. Modulators of this protein are used to treat or prevent  
CC conditions associated with abnormal levels of inflammation, or activity  
CC of IL-1 or its receptor complex

XX SQ Sequence 5362 BP; 1955 A; 898 C; 989 G; 1519 T; 0 U; 1 Other;  
Query Match 61.8%; Score 20.4; DB 2; Length 5362;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTCTGCTTGCTT 31  
||||| ||||| ||||| ||||| |||||  
Db 5214 CCATCTATGAGTTAATTCCTGTTGCTT 5185

RESULT 20  
ADF41968  
ID ADF41968 standard; DNA; 1006 BP.

XX AC ADF41968;

XX DT 12-FEB-2004 (first entry)

XX DE Alcaligenes faecalis PtxD DNA.

XX KW phosphite dehydrogenase; phosphate; reduced nucleotide cofactor; NADH;  
XX KW NADPH; chiral; carbohydrate; PtxD; ds; gene.

XX OS Alcaligenes faecalis.

XX PN WO2003072726-A2.

XX PD 04-SEP-2003.

XX PF 21-FEB-2003; 2003WO-US005360.

XX PR 22-FEB-2002; 2002US-0359091P.

XX PA (BIOT-) BIOTECHNOLOGY RES & DEV CORP.

XX PA (UYIL-) UNIV ILLINOIS URBANA-CHAMPAIGN.

XX PI Metcalf W, Van Der Donk WA, Vrtis JM, White AK, Garcia Costas AM;

XX DR WPI; 2003-731607/69.

XX DR P-PSDB; ADF41967.

XX PT New phosphite dehydrogenase enzyme capable of converting phosphite to  
XX PT phosphate, useful for regenerating cofactors, e.g. NAD(P)H, or for  
XX PT synthesizing chiral compounds, complex carbohydrates or isotopically-  
XX PT labeled compounds.

XX PS Claim 39; Page 51-52; 90pp; English.

XX CC The invention relates to a novel purified enzyme capable of converting  
XX CC phosphite to phosphate comprising a phosphite dehydrogenase produced by a  
XX CC recombinant process or isolated from a natural source. The enzyme and,  
XX CC methods of the invention may be useful in regenerating reduced nucleotide  
XX CC cofactors, such as NADH and NADPH, or in synthesizing chiral compounds,  
XX CC complex carbohydrates or isotopically-labelled compounds. The current  
XX CC sequence is that of the Alcaligenes faecalis PtxD DNA of the invention.

XX SQ Sequence 1006 BP; 226 A; 268 C; 289 G; 223 T; 0 U; 0 Other;

Query Match 61.2%; Score 20.2; DB 10; Length 1006;  
Best Local Similarity 88.0%; Pred. No. 1.3e+02;  
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 TTACGCTAGTTCCTGCTTGCTTAC 33  
||||| ||||| ||||| ||||| |||||

Db 278 TTACGATGTTCTGCTGATTGCTTAC 302

RESULT 21

ABA18630

ID ABA18630 standard; DNA; 8668 BP.

XX AC ABA18630;

XX DT 23-JAN-2002 (first entry)

XX DE Human nervous system related polynucleotide SEQ ID NO 10961.

XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
XX KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
XX KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;  
XX KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
XX KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
XX KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
XX KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

XX OS Homo sapiens.

XX PN WO200159063-A2.

XX PD 16-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001334.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214866P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 14-JUL-2000; 2000US-0217496P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 14-AUG-2000; 2000US-0225758P.

XX PR 18-AUG-2000; 2000US-0225759P.

XX PR 22-AUG-2000; 2000US-0226279P.

XX PR 22-AUG-2000; 2000US-0226681P.

XX PR 22-AUG-2000; 2000US-0226688P.

XX PR 22-AUG-2000; 2000US-0227182P.

XX PR 23-AUG-2000; 2000US-0227009P.

XX PR 30-AUG-2000; 2000US-0228924P.

XX PR 01-SEP-2000; 2000US-0229287P.

XX PR 01-SEP-2000; 2000US-0229343P.

XX PR 01-SEP-2000; 2000US-0229345P.

XX PR 05-SEP-2000; 2000US-0229509P.

XX PR 06-SEP-2000; 2000US-0229513P.

XX PR 06-SEP-2000; 2000US-0230437P.

XX PR 08-SEP-2000; 2000US-0231242P.

XX PR 08-SEP-2000; 2000US-0231243P.

XX PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239933P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 20-OCT-2000; 2000US-0242221P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0244674P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 01-DEC-2000; 2000US-0251160P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI WPT; 2001-541565/60.  
 XX  
 DR Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
 XX useful for preventing, diagnosing and/or treating nervous system cancers  
 XX and metastases.  
 XX Disclosure; SEQ ID NO 10961; 1701pp + Sequence Listing; English.  
 PS  
 CC The invention relates to novel genes (AB11004-AB21534) and proteins  
 CC (AB114678-AB18001) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, diabetes mellitus, Crohn's disease,  
 CC anaemia, autoimmune thyroiditis, allergies, autoimmune haemolytic  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 8668 BP; 2790 A; 1549 C; 1854 G; 2475 T; 0 U; 0 Other;  
 Query Match 61.2%; Score 20.2; DB 5; Length 8668;  
 Best Local Similarity 88.0%; Pred. No. 1.8e+02;  
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 8 ATTACGCTAGTCTCTGTTGCTTA 32  
 |||||  
 Db 4018 ATTATGCTAGTCTCTGTTGCTTA 4042  
 RESULT 22  
 ABD32649/c  
 ID ABD32649 standard; DNA; 102790 BP.  
 XX  
 AC ABD32649;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Human cancer-associated genomic DNA HD13-036.  
 XX  
 KW Human; ds; cancer-associated protein; gene; cytostatic; cancer;  
 KW Leukaemia; lymphoma; CAP.  
 XX  
 OS Homo sapiens.  
 XX



XX Brogginì M, D'incalci M;  
XX WPI; 2003-221715/21.  
XX New oncosuppressive polypeptide, useful for preparing a medicament for  
XX treating tumors.  
XX Claim 3; Page 40-41; 42pp; English.  
XX The invention relates to oncosuppressive polypeptides and encoding  
XX polynucleotides. The oncosuppressive gene is involved in apoptotic  
XX process and is regulated by p53 and p73. The oncosuppressive  
XX polynucleotides are useful for preparing a medicament for treating  
XX tumour. Sequences ABZ58998-9004 represent DNA fragments of the human  
XX oncosuppressive gene (DRAGO), having promoter activity responsive to p53  
XX and p73  
XX Sequence 1203 BP; 305 A; 237 C; 274 G; 387 T; 0 U; 0 Other;  
SQ

Query Match 60.6%; Score 20; DB 8; Length 1203;  
Best Local Similarity 82.1%; Pred. No. 1.7e+02;  
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 6 TTATTACGCTAGTTCCTGGTTTGCTTAC 33  
Db 779 TTATACTGCTAGTTCCTGGTTTGCTTAC 806

RESULT 25  
ABZ59003  
ID ABZ59003 standard; DNA; 1637 BP.  
XX  
XX ABZ59003;  
XX  
XX 28-APR-2003 (first entry)  
XX Human oncosuppressive gene (DRAGO) fragment.  
XX  
XX Oncosuppressive; apoptotic; p53; p73; cytostatic; gene therapy; tumour;  
XX DRAGO; human; promoter; gene; ds.  
XX Homo sapiens.  
XX WO2003006498-A2.  
XX  
XX 23-JAN-2003.  
XX  
XX 09-JUL-2002; 2002WO-EP007625.  
XX  
XX 10-JUL-2001; 2001IT-MI001465.  
XX  
XX (NOVU-) NOVUSPHARMA SPA.  
XX  
XX Brogginì M, D'incalci M;  
XX WPI; 2003-221715/21.  
XX New oncosuppressive polypeptide, useful for preparing a medicament for  
XX treating tumors.  
XX Claim 3; Page 41-42; 42pp; English.  
XX The invention relates to oncosuppressive polypeptides and encoding  
XX polynucleotides. The oncosuppressive gene is involved in apoptotic  
XX process and is regulated by p53 and p73. The oncosuppressive  
XX polynucleotides are useful for preparing a medicament for treating  
XX tumour. Sequences ABZ58998-9004 represent DNA fragments of the human  
XX oncosuppressive gene (DRAGO), having promoter activity responsive to p53  
XX and p73  
XX Sequence 1637 BP; 413 A; 329 C; 381 G; 514 T; 0 U; 0 Other;  
SQ

Query Match 60.6%; Score 20; DB 8; Length 1637;  
Best Local Similarity 82.1%; Pred. No. 1.7e+02;  
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 6 TTATTACGCTAGTTCCTGGTTTGCTTAC 33  
Db 779 TTATACTGCTAGTTCCTGGTTTGCTTAC 806

RESULT 26  
ABZ58998  
ID ABZ58998 standard; DNA; 2464 BP.  
XX  
XX ABZ58998;  
XX  
XX 28-APR-2003 (first entry)  
XX Human oncosuppressive gene (DRAGO) fragment.  
XX  
XX Oncosuppressive; apoptotic; p53; p73; cytostatic; gene therapy; tumour;  
XX DRAGO; human; promoter; gene; ds.  
XX Homo sapiens.  
XX WO2003006498-A2.  
XX  
XX 23-JAN-2003.  
XX  
XX 09-JUL-2002; 2002WO-EP007625.  
XX  
XX 10-JUL-2001; 2001IT-MI001465.  
XX  
XX (NOVU-) NOVUSPHARMA SPA.  
XX  
XX Brogginì M, D'incalci M;  
XX WPI; 2003-221715/21.  
XX New oncosuppressive polypeptide, useful for preparing a medicament for  
XX treating tumors.  
XX Claim 3; Page 39-40; 42pp; English.  
XX The invention relates to oncosuppressive polypeptides and encoding  
XX polynucleotides. The oncosuppressive gene is involved in apoptotic  
XX process and is regulated by p53 and p73. The oncosuppressive  
XX polynucleotides are useful for preparing a medicament for treating  
XX tumour. Sequences ABZ58998-9004 represent DNA fragments of the human  
XX oncosuppressive gene (DRAGO), having promoter activity responsive to p53  
XX and p73  
XX Sequence 2464 BP; 611 A; 530 C; 551 G; 772 T; 0 U; 0 Other;  
SQ

Query Match 60.6%; Score 20; DB 8; Length 2464;  
Best Local Similarity 82.1%; Pred. No. 1.8e+02;  
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 6 TTATTACGCTAGTTCCTGGTTTGCTTAC 33  
Db 779 TTATACTGCTAGTTCCTGGTTTGCTTAC 806

RESULT 27  
ABZ58996  
ID ABZ58996 standard; DNA; 3761 BP.  
XX  
XX ABZ58996;  
XX  
XX 28-APR-2003 (first entry)  
XX Human oncosuppressive gene (DRAGO) fragment.  
XX  
XX Oncosuppressive; apoptotic; p53; p73; cytostatic; gene therapy; tumour;  
XX

```
KW DRAGO; human; gene; ds.
XX
OS Homo sapiens.
XX
XX WO2003006498-A2.
XX
XX 23-JAN-2003.
XX
XX 09-JUL-2002; 2002WO-EP007625.
XX
XX 10-JUL-2001; 2001IT-MI001465.
XX
XX (NOVU-) NOVUSPHARMA SPA.
XX
XX Brogginini M, D'incalci M;
XX
XX WPI; 2003-221715/21.
XX
XX New oncosuppressive polypeptide, useful for preparing a medicament for
XX treating tumors.
XX
PS Claim 3; Page 37-38; 42pp; English.
XX
CC The invention relates to oncosuppressive polypeptides and encoding
CC polynucleotides. The oncosuppressive gene is involved in apoptotic
CC process and is regulated by p53 and p73. The oncosuppressive
CC polynucleotides are useful for preparing a medicament for treating
CC tumour. The present sequence represents a human oncosuppressive gene
CC (DRAGO) fragment.
XX
XX Query Match 60.6%; Score 20; DB 8; Length 3761;
XX Best Local Similarity 82.1%; Pred. No. 1.9e+02;
XX Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
QY 6 TTATTACGCTAGTCTCTGTTGCTTAC 33
DB 1917 TTATACTGCTAGTCTCTGTTGCTTAC 1944

RESULT 28
ABZ58995
ID ABZ58995 standard; DNA; 10735 BP.
XX
XX AC ABZ58995;
XX
XX 28-APR-2003 (first entry)
XX
XX Human oncosuppressive gene (DRAGO) fragment.
XX
XX Oncosuppressive; apoptotic; p53; p73; cytostatic; gene therapy; tumour;
XX DRAGO; human; gene; ds.
XX
XX Homo sapiens.
XX
XX WO2003006498-A2.
XX
XX 23-JAN-2003.
XX
XX 09-JUL-2002; 2002WO-EP007625.
XX
XX 10-JUL-2001; 2001IT-MI001465.
XX
XX (NOVU-) NOVUSPHARMA SPA.
XX
XX Brogginini M, D'incalci M;
XX
XX WPI; 2003-221715/21.
XX
XX New oncosuppressive polypeptide, useful for preparing a medicament for
XX treating tumors.
XX

KW DRAGO; human; gene; ds.
XX
OS Homo sapiens.
XX
XX WO2003006498-A2.
XX
XX 23-JAN-2003.
XX
XX 09-JUL-2002; 2002WO-EP007625.
XX
XX 10-JUL-2001; 2001IT-MI001465.
XX
XX (NOVU-) NOVUSPHARMA SPA.
XX
XX Brogginini M, D'incalci M;
XX
XX WPI; 2003-221715/21.
XX
XX New oncosuppressive polypeptide, useful for preparing a medicament for
XX treating tumors.
XX

PS Claim 3; Page 34-37; 42pp; English.
XX
CC The invention relates to oncosuppressive polypeptides and encoding
CC polynucleotides. The oncosuppressive gene is involved in apoptotic
CC process and is regulated by p53 and p73. The oncosuppressive
CC polynucleotides are useful for preparing a medicament for treating
CC tumour. The present sequence represents a human oncosuppressive gene
CC (DRAGO) fragment, located upstream of the first exon
XX
XX Query Match 60.6%; Score 20; DB 8; Length 10735;
XX Best Local Similarity 82.1%; Pred. No. 2.2e+02;
XX Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
QY 6 TTATTACGCTAGTCTCTGTTGCTTAC 33
DB 8891 TTATACTGCTAGTCTCTGTTGCTTAC 8918

RESULT 29
ADQ97643
ID ADQ97643 standard; DNA; 100137 BP.
XX
XX AC ADQ97643;
XX
XX 07-OCT-2004 (first entry)
XX
XX Human cancer associated sequence HD10-020, SEQ ID 620.
XX
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
XX Homo sapiens.
XX
XX WO2004060304-A2.
XX
XX 22-JUL-2004.
XX
XX 22-DEC-2003; 2003WO-US041389.
XX
XX 27-DEC-2002; 2002US-00330773.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Malandro MS;
XX
XX WPI; 2004-543781/52.
XX
XX New isolated cancer associated nucleic acids comprising at least 10
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX cancers such as leukemia and lymphoma.
XX
XX Claim 1; SEQ ID NO 620; 199pp; English.
XX
XX The present invention relates to cancer associated sequences (ADQ97025-
XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Query Match 60.6%; Score 20; DB 12; Length 100137;
XX Best Local Similarity 82.1%; Pred. No. 2.9e+02;
XX Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
QY 5 TTTATTACGCTAGTCTCTGTTGCTTAC 32
DB 66010 TTTATTACGCTAGTCTCTGTTGCTTAC 66037

RESULT 30
```

```
AAQ51631
ID AAQ51631 standard; DNA; 369 BP.
XX
AC AAQ51631;
XX
DT 10-MAR-2003 (revised)
DT 23-JUN-1994 (first entry)
XX
DE Anti-lysozyme VH gene.
XX
KW Antibody; lysozyme; primer; amplification; mutation;
KW polymerase chain reaction; PCR; ds.
XX
OS Unidentified.
XX
PN JP05276946-A.
XX
PD 26-OCT-1993.
XX
PF 14-OCT-1991; 91JP-00291964.
XX
PR 14-OCT-1991; 91JP-00291964.
XX
PA (FUJI-) FUJITA GAKUEN GH.
PA (TOYJ-) TOSOH CORP.
XX
WPI; 1993-373583/47.
DR P-PSDB; AAR43827.
XX
DNA prepn. using PCR techniques - where PCR is performed using primer
PT mixt. of which one of the oligo-nucleotide primers has different
PT nucleotide sequence and hybridise on identical region.
XX
PS Disclosure; Fig 1; 11pp; Japanese.
XX
DNA was isolated from anti-lysozyme hybridoma Hyb Cl. VH gene was
CC isolated and the nucleotide sequence determined. PCR was performed using
CC primers VHXho, VHECOBND and PalBECO. Mutated primers are LyHX40, LyHN24,
CC LyHNI140 and LyHS46. (Updated on 10-MAR-2003 to add missing OS field.)
XX
SQ Sequence 369 BP; 96 A; 79 C; 104 G; 90 T; 0 U; 0 Other;

Query Match 60.0%; Score 19.8; DB 2; Length 369;
Best Local Similarity 77.4%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGGTTTGCTTAC 33
Db 306 CAGTGATTACGACCTTGCCCTGGTTTGCTTAC 336

RESULT 31
ABN70767/c
ID ABN70767 standard; DNA; 966 BP.
XX
AC ABN70767;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polymnucleotide SEQ ID NO 9447.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus agalactiae.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
PR 29-OCT-2001; 2001WO-GB004789.
XX

AAQ51631
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Masignani V, Margarit Y Rosl, Grandi G, Fraser C;
PI Tettelin H;
XX
WPI; 2002-352536/38.
DR P-PSDB; ABP30136.
XX
New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
PS Claim 7; Page 4070-4071; 4525pp; English.
XX
The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 966 BP; 278 A; 161 C; 205 G; 322 T; 0 U; 0 Other;

Query Match 60.0%; Score 19.8; DB 6; Length 966;
Best Local Similarity 77.4%; Pred. No. 2e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGGTTTGCTTAC 33
Db 63 CACITTTTCGCTAGTTCCTGGTTTGCTTAC 33

RESULT 32
ABN66547/c
ID ABN66547 standard; DNA; 978 BP.
XX
AC ABN66547;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polymnucleotide SEQ ID NO 1007.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus agalactiae.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
```





```

XX  Jiang B, Hu W, Lemieux S, Roemer T;
PI  WPI; 2004-594200/57.
DR  P-PSDB; ADR86302.
XX
XX  New purified or isolated Aspergillus fumigatus nucleic acid molecule
PT  encoding a gene product, useful for diagnosing and/or treating invasive
PT  fungal infections, such as Farmer's lung disease.
XX
XX  Claim 3; SEQ ID NO 352; 164pp; English.
XX
XX  The present invention relates to Aspergillus fumigatus genes that are
CC  essential and are potential targets for drug screening. The methods and
CC  compositions of the present invention are useful for diagnosing and/or
CC  treating invasive Aspergillus fumigatus infection, including the allergic
CC  forms of the disease, such as Farmer's lung disease. They can also be
CC  used in various drug discovery purposes, such as expression of the
CC  recombinant protein, hybridization assay and construction of nucleic acid
CC  arrays. The present sequence represents an Aspergillus fumigatus
CC  essential gene full length genomic sequence, used during diagnosis and
CC  drug development in the invention. These genes share a high degree of
CC  sequence conservation with known essential genes of candida albicans. The
CC  sequence data for this patent is not represented in the printed
CC  specification, but was obtained in electronic format from WIPO.
XX
XX  Sequence 7179 BP; 1813 A; 1736 C; 1855 G; 1775 T; 0 U; 0 Other;
SQ
XX  Query Match 60.0%; Score 19.8; DB 13; Length 7179;
XX  Best Local Similarity 77.4%; Pred. No. 2.5e+02;
XX  Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX  3 CATTATTACGCTAGTCTCTGCTTTGCTTAC 33
XX  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX  5717 CATTGAGCCGCTTGTTCAGGTTTGCTGAC 5687
XX
XX  RESULT 35
XX  AAD47150
XX  ID AAD47150 standard; DNA; 31814 BP.
XX
XX  AC AAD47150;
XX
XX  24-FEB-2003 (first entry)
XX
XX  Human Ras-like protein encoding gene.
XX
XX  Human; Ras-like protein; inflammation; cell proliferation; apoptosis;
XX  immunodeficiency; neurodegenerative; Alzheimer's disease; cirrhosis;
XX  Parkinson's disease; wasting disease; cachexia; myocardial infarction;
XX  osteoporosis; atherosclerosis; glomerulonephritis; Crohn's disease;
XX  irritable bowel syndrome; multiple sclerosis; osteoarthritis; allergy;
XX  pancreatitis; autoimmune thyroiditis; rheumatoid arthritis; osteopathic;
XX  Sjogren's syndrome; infection; transgenic; gene therapy; neutropenic;
XX  neuroprotective; vasotrophic; cytostatic; dermatological; nephrotropic;
XX  ophthalmological; tranquilizer; cancer; stroke; Grave's disease; AIDS;
XX  asthma; anaemia; chromosome 3; gene; ds.
XX
XX  Homo sapiens.
XX
XX  Key Location/Qualifiers
XX  variation replace(1193,T)
XX  /tag= a
XX  /standard_name= "Single nucleotide polymorphism"
XX
XX  CDS 2305..28816
XX  /tag= b
XX
XX  FT /product= "Human Ras-like protein"
XX  exon 2305..2494
XX  /tag= c
XX  FT intron 2495..28679
XX  /tag= d
XX  FT variation replace(9276, -)
XX  /tag= e

```

```

FT  exon /standard_name= "Single nucleotide polymorphism"
FT  28680..28813
FT  /tag= f
XX
XX  WO200277190-A2.
XX
XX  03-OCT-2002.
XX
XX  27-MAR-2002; 2002WO-US009323.
XX
XX  27-MAR-2001; 2001US-00817182.
XX  (PEKE ) PE CORP NY.
XX
XX  Merkulov GV, Gong F, Di Francesco V, Beasley EM;
PI  WPI; 2003-018912/01.
DR  P-PSDB; AAE29095.
XX
XX  New isolated human Ras-like protein polypeptides, useful for diagnosing,
PT  treating or preventing inflammation and disorders associated with cell
PT  proliferation and apoptosis, e.g. AIDS, cancer, allergies, anemia, asthma
PT  or stroke.
XX
XX  Claim 22; Page 69-77; 78pp; English.
XX
XX  The invention relates to human Ras-like protein and its corresponding
XX  nucleic acid. The Ras-like protein and DNA is useful in the development
XX  of human therapeutics and diagnostic compositions. They are useful in the
XX  diagnosis, prevention and treatment of inflammation and disorders
XX  associated with cell proliferation and apoptosis, e.g. AIDS and other
XX  infectious or genetic immunodeficiencies, neurodegenerative disease e.g.
XX  Alzheimer's disease, Parkinson's disease, wasting diseases e.g. cachexia,
XX  ischaemic injuries e.g. myocardial infarction, stroke or reperfusion
XX  injury, toxin-induced diseases such as alcohol-induced liver damage or
XX  cirrhosis, osteoporosis or cancer. They are also used to treat disorders
XX  associated with inflammation including allergies, atopic dermatitis,
XX  atherosclerosis, asthma, anaemia, Crohn's disease, diabetes mellitus,
XX  Grave's disease, glomerulonephritis, gout, irritable bowel syndrome,
XX  lupus erythematosus, multiple sclerosis, osteoarthritis, pancreatitis,
XX  autoimmune thyroiditis, rheumatoid arthritis, Sjogren's syndrome,
XX  uveitis, trauma, or viral, bacterial, fungal, parasitic, protozoal or
XX  helminthic infections. The antibodies of the invention are useful in
XX  pharmacogenomic analysis or for tissue typing. The transgenic animals are
XX  useful for studying the function of a Ras-like protein, and identifying
XX  and evaluating modulators of its activity. The DNA of the invention is
XX  used in gene therapy. The present sequence is human Ras-like protein
XX  encoding gene. This gene is located at chromosome 3
XX
XX  Sequence 31814 BP; 6426 A; 6968 C; 7462 G; 7934 T; 0 U; 3024 Other;
SQ
XX  Query Match 60.0%; Score 19.8; DB 10; Length 31814;
XX  Best Local Similarity 77.4%; Pred. No. 3.1e+02;
XX  Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX  Qy 1 CCCATTATTACGCTAGTCTCTGCTTTGCTT 31
XX  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX  Db 13561 CCCATTATCCAGCTACCTCTTGCTTCTT 13591
XX
XX  RESULT 36
XX  ACN44168
XX  ID ACN44168 standard; DNA; 39121 BP.
XX
XX  AC ACN44168;
XX
XX  18-NOV-2004 (first entry)
XX
XX  Mouse genomic sequence MCG15397.
XX
XX  Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
XX  Mus musculus.

```

XX WO2003073826-A2.  
XX 12-SEP-2003.  
XX 28-FEB-2003; 2003WO-US006235.  
XX 01-MAR-2002; 2002US-00087192.  
XX (SAGR-) SAGRES DISCOVERY.  
XX Morris DW;  
XX WPI; 2003-328604/31.  
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
PT comprises a nucleotide sequence.  
XX Claim 1; SEQ ID NO 481; Opp; English.  
XX The present invention relates to novel DNA and protein sequences which  
CC are associated with carcinomas. The sequences are useful for: (i) for  
CC screening drug candidates; (ii) for screening of bioactive agent capable  
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
CC a bioactive agent capable of modulating the activity of CAP; (iv) for  
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
CC determining Carcinoma Associated (CA) gene copy number. In addition, the  
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of  
CC carcinoma including lymphoma. The present sequence is one such CA coding  
CC sequence. Note: This patent is an equivalent to basic patent  
CC US2002182586A1, for which no sequence data was published  
XX  
XX Sequence 39121 BP; 9718 A; 6795 C; 7194 G; 11535 T; 0 U; 3879 Other;  
Query Match 60.0%; Score 19.8; DB 11; Length 39121;  
Best Local Similarity 77.4%; Pred. No. 3.1e+02;  
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 3 CATTATTACGCTAGTCTCTGGTTGCTTAC 33  
Db 20031 CATTATGACTCTGATTCTGGTTGCTTCC 20061  
RESULT 37  
ABN71527\_00/c  
WP Sequence split into 22 fragments LOCUS ABN71527 Accession Abn71527  
WP Fragment Name Begin End  
WP ABN71527\_00 1 110000  
WP ABN71527\_01 100001 210000  
WP ABN71527\_02 200001 310000  
WP ABN71527\_03 300001 410000  
WP ABN71527\_04 400001 510000  
WP ABN71527\_05 500001 610000  
WP ABN71527\_06 600001 710000  
WP ABN71527\_07 700001 810000  
WP ABN71527\_08 800001 910000  
WP ABN71527\_09 900001 1010000  
WP ABN71527\_10 1000001 1110000  
WP ABN71527\_11 1100001 1210000  
WP ABN71527\_12 1200001 1310000  
WP ABN71527\_13 1300001 1410000  
WP ABN71527\_14 1400001 1510000  
WP ABN71527\_15 1500001 1610000  
WP ABN71527\_16 1600001 1710000  
WP ABN71527\_17 1700001 1810000  
WP ABN71527\_18 1800001 1910000  
WP ABN71527\_19 1900001 2010000  
WP ABN71527\_20 2000001 2110000  
WP ABN71527\_21 2100001 2155561  
ID ABN71527 standard; DNA; 2155561 BP.

XX ABN71527;  
XX 02-JUL-2002 (first entry)  
XX Streptococcus polynucleotide SEQ ID NO 10967.  
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;  
KW antinflammatory; infection; vaccine; meningitis; gene therapy; ds.  
XX Streptococcus sp.  
XX WO200234771-A2.  
XX 02-MAY-2002.  
XX 29-OCT-2001; 2001WO-GB004789.  
XX 27-OCT-2000; 2000GB-00026333.  
XX 24-NOV-2000; 2000GB-00028727.  
XX 07-MAR-2001; 2001GB-00005640.  
XX (CHIR-) CHIRON SPA.  
XX (GENO-) INST GENOMIC RES.  
XX Telford J, Maignani V, Margarit Y Rosi, Grandi G, Fraser C;  
XX Tettelin H;  
XX WPI; 2002-352536/38.  
XX New Streptococcus protein for the treatment or prevention of infection or  
PT disease caused by Streptococcus bacteria, such as meningitis, and for  
PT detecting a compound that binds to the protein.  
XX Claim 8; Page 4196-4488; 4525pp; English.  
XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins  
XX Sequence 2155561 BP; 695741A; 385320C; 382958G; 691542T; 0U; 0Other;  
Query Match 60.0%; Score 19.8; DB 6; Length 110000;  
Best Local Similarity 77.4%; Pred. No. 3.6e+02;  
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 3 CATTATTACGCTAGTCTCTGGTTGCTTAC 33  
Db 1644 CACTTTTTCGCTAGTCTCTTCTTCTTGTGATGAC 1614  
RESULT 38  
ACN43998\_0  
WP Sequence split into 7 fragments LOCUS ACN43998 Accession ACN43998  
WP Fragment Name Begin End  
WP ACN43998\_0 1 110000  
WP ACN43998\_1 100001 210000  
WP ACN43998\_2 200001 310000

```
WP ACN43998_3 300001 410000
WP ACN43998_4 400001 510000
WP ACN43998_5 500001 610000
WP ACN43998_6 600001 653122
ID ACN43998 standard; DNA; 653122 BP.
XX
AC ACN43998;
XX
XX 18-NOV-2004 (first entry)
XX Human genomic sequence hCG1778483.
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
XX Homo sapiens.
XX WO2003073826-A2.
XX
XX 12-SEP-2003.
XX
XX 28-FEB-2003; 2003WO-US006235.
XX
XX 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW;
XX
XX WPI; 2003-328604/31.
XX
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.
XX
XX Claim 1; SEQ ID NO 226; Opp; English.
XX
XX The present invention relates to novel DNA and protein sequences which
XX are associated with carcinomas. The sequences are useful for: (i) for
XX screening drug candidates; (ii) for screening of bioactive agent capable
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX a bioactive agent capable of modulating the activity of CAP; (iv) for
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX carcinoma including lymphoma. The present sequence is one such CA coding
XX sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published
XX
XX Sequence 653122 BP; 180452A; 129260C; 137452G; 202235T; 0U; 37230Other;
XX
XX Query Match 60.0%; Score 19.8; DB 11; Length 110000;
XX Best Local Similarity 77.4%; Pred. No. 3.6e+02;
XX Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX 1 CCCATTATTACGCTAGTTCCTGGTTTGCTT 31
XX |||||
XX 42506 CACATTATTACTTTGGTTTGGTATGCTT 42536
XX
XX RESULT 39
XX ADQ18503
XX ID ADQ18503 standard; DNA; 310 BP.
XX
XX AC ADQ18503;
XX
XX 26-AUG-2004 (first entry)
XX
XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 1322.
XX
XX soft tissue sarcoma; cytotstatic; gene therapy; vaccine; screening; human;
XX ds.
XX
```

```
XX Homo sapiens.
XX WO2004048938-A2.
XX 10-JUN-2004.
XX 26-NOV-2003; 2003WO-US038193.
XX 26-NOV-2002; 2002US-0429739P.
XX (PROT-) PROTEIN DESIGN LABS INC.
XX Aziz N, Ginsburg WM, Zlotnik A;
XX WPI; 2004-441208/41.
XX Early detection of soft tissue sarcoma comprises determining expression
XX of a gene in a first soft tissue sample and a normal soft tissue sample
XX and comparing the gene expression, also useful in treating soft tissue
XX sarcoma.
XX
XX Example 2; SEQ ID NO 1322; 210pp; English.
XX
XX The invention relates to a novel method for detecting soft tissue sarcoma
XX which comprises obtaining a first soft tissue sample from an individual
XX and a normal soft tissue sample from the same or different individual,
XX determining the expression of a gene in both samples and comparing the
XX expression of the gene in both soft tissue samples, where a higher level
XX of protein expression in the first soft tissue sample indicates the
XX presence of soft tissue sarcoma. The method of the invention has
XX cytostatic applications and may be useful for detecting soft tissue
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX acid sequences may be useful in diagnostic and screening applications.
XX The current sequence is that of a human soft tissue sarcoma-upregulated
XX DNA of the invention. The current sequence is not shown within the
XX specification per se but was submitted in CD format by the inventor.
XX
XX Sequence 310 BP; 72 A; 84 C; 72 G; 82 T; 0 U; 0 Other;
XX
XX Query Match 59.4%; Score 19.6; DB 12; Length 310;
XX Best Local Similarity 84.6%; Pred. No. 2.1e+02;
XX Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX 2 CCATTATTACGCTAGTTCCTGGTTT 27
XX |||||
XX 34 CCATTATTACCTAGTCTGGTTT 59
XX
XX RESULT 40
XX AAH94452
XX ID AAH94452 standard; cDNA; 617 BP.
XX
XX AC AAH94452;
XX
XX 05-OCT-2001 (first entry)
XX
XX Human foetal cDNA, SEQ ID NO: 1139.
XX
XX Human; foetal protein; cytotstatic; immunosuppressive; immunostimulant;
XX neutropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
XX gene therapy; antitense therapy; cancer; immune disorder;
XX growth disorder; osteoporosis; thrombolytic disorder;
XX nervous system disorder; inflammation; expressed sequence tag; EST; ss.
XX
XX Homo sapiens.
XX
XX WO200155339-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US002723.
XX
```

```
PR 25-JAN-2000; 2000US-00491404.
PR 15-SEP-2000; 2000US-00663870.
PR 06-NOV-2000; 2000US-00707351.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Yeung G, Ford JE, Boyle BJ, Arterburn MC, Dmanac RA, Tang YT;
XX Liu C, Asundi V, Zhou P, Werhman T;
XX
XX WPI; 2001-465571/50.
XX P-PSDB; AAM06777.
XX
XX Novel fetal proteins useful for the treatment and diagnosis of diseases
XX PT associated with dysfunction of the protein e.g. cancers, immune
XX PT disorders, growth disorders, thrombolytic disorders, nervous system
XX PT disorders and inflammation.
XX
XX Example 3; Page 638-639; 715pp; English.
XX
XX The invention relates to novel foetal polypeptides encoded by
XX CC polynucleotides comprising one of 477 sequences fully defined in the
XX CC specification. The foetal polynucleotides and polypeptides are useful in
XX CC the treatment and diagnosis of diseases such as cancers, immune
XX CC disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders,
XX CC nervous system disorders and inflammation. The present sequence is a full
XX CC length cDNA which was assembled using expressed sequence tags (ESTs)
XX CC found to be expressed in human foetal tissue cDNA libraries as seeds
XX
XX SQ Sequence 617 BP; 192 A; 124 C; 96 G; 205 T; 0 U; 0 Other;
XX
XX Query Match 58.8%; Score 19.4; DB 5; Length 617;
XX Best Local Similarity 79.3%; Pred. No. 2.7e+02;
XX Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 2 CCATTATTACGCTAGTTCCTCGTTTGCT 30
XX | | | | | | | | | | | | | | | | | |
XX Db 575 CAATCTAATGTGCTAGTTCCTCGTTTGCT 603
XX
XX Search completed: June 4, 2005, 13:53:56
XX Job time : 205.333 secs
```

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 13:25:28 ; Search time 65 Seconds  
(without alignments)  
830.725 Million cell updates/sec

Title: US-10-089-452-26

Perfect score: 33

Sequence: 1 cccattattacgtagtctctggttgcttac 33

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21.4	64.8	58812	4	US-09-949-016-17197
2	21.4	64.8	95566	4	US-09-949-016-11877
3	20.8	63.0	3501	4	US-09-809-665A-70
4	20.4	61.8	152331	3	US-09-128-155-16
5	20.4	61.8	176373	3	US-09-128-155-17
6	19.8	60.0	2710	4	US-08-956-171E-421
7	19.8	60.0	2710	4	US-08-781-986A-421
8	19.6	59.4	13445	4	US-09-949-016-13512
9	19.4	58.8	601	4	US-09-949-016-107885
10	19.4	58.8	115388	4	US-09-949-016-14981
11	19.4	58.8	390890	4	US-09-949-016-14982
12	19.4	58.8	390890	4	US-09-949-016-14720
13	19.2	58.2	660	4	US-09-543-681A-2150
14	19.2	58.2	65518	4	US-09-949-016-12421
15	19.2	58.2	65518	4	US-09-949-016-14101
16	19.2	58.2	67386	4	US-09-949-016-16519
17	19	57.6	601	4	US-09-949-016-150746
18	19	57.6	118143	4	US-09-949-016-17196
19	18.8	57.0	599	3	US-09-242-901-4
20	18.8	57.0	2938	3	US-09-072-917A-8
21	18.8	57.0	56976	4	US-09-949-016-17486
22	18.8	57.0	103674	4	US-09-949-016-12033
23	18.8	57.0	253345	4	US-09-949-016-12656
24	18.8	57.0	253364	4	US-09-949-016-13639
25	18.8	57.0	263693	4	US-09-949-016-12386
26	18.8	57.0	263694	4	US-09-949-016-16915
27	18.6	56.4	409	1	US-08-400-256-37

C 28	18.6	56.4	409	3	US-08-975-365-37	Sequence 37, Appl
C 29	18.6	56.4	415	1	US-08-400-256-16	Sequence 16, Appl
C 30	18.6	56.4	415	1	US-08-400-256-22	Sequence 22, Appl
C 31	18.6	56.4	415	1	US-08-400-256-25	Sequence 25, Appl
C 32	18.6	56.4	415	1	US-08-400-256-28	Sequence 28, Appl
C 33	18.6	56.4	415	1	US-08-400-256-31	Sequence 31, Appl
C 34	18.6	56.4	415	3	US-08-975-365-16	Sequence 16, Appl
C 35	18.6	56.4	415	3	US-08-975-365-22	Sequence 22, Appl
C 36	18.6	56.4	415	3	US-08-975-365-25	Sequence 25, Appl
C 37	18.6	56.4	415	3	US-08-975-365-28	Sequence 28, Appl
C 38	18.6	56.4	415	3	US-08-975-365-31	Sequence 31, Appl
C 39	18.6	56.4	601	4	US-09-949-016-206088	Sequence 206088,
C 40	18.6	56.4	601	4	US-09-949-016-206125	Sequence 206125,
C 41	18.6	56.4	91559	4	US-09-949-016-12581	Sequence 12581, A
C 42	18.6	56.4	91559	4	US-09-949-016-13701	Sequence 13701, A
C 43	18.6	56.4	113379	4	US-09-949-016-17561	Sequence 17561, A
C 44	18.6	56.4	113379	4	US-09-949-016-17562	Sequence 17562, A
C 45	18.4	55.8	555	4	US-09-248-796A-6339	Sequence 6339, Ap

#### ALIGNMENTS

##### RESULT 1

US-09-949-016-17197  
; Sequence 17197, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17197  
; LENGTH: 58812  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-949-016-17197

Query Match 64.8%; Score 21.4; DB 4; Length 58812;  
Best Local Similarity 80.6%; Pred. No. 20;  
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCCATTATTACGCTAGTTCCTGGTTGCTT 31

Db 34016 CCCATTAAACTCATTTTCATGTTGCTT 34046

##### RESULT 2

US-09-949-016-11877  
; Sequence 11877, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11877  
; LENGTH: 95566  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-11877

Query Match 64.8%; Score 21.4; DB 4; Length 95566;  
Best Local Similarity 80.6%; Pred. No. 22;  
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCATTATTACGCTAGTTCCTGGTTTGCTT 31  
DB 70775 CCATTAAACTATTCATGTTTGCTT 70805

RESULT 3  
US-09-809-665A-70  
; Sequence 70, Application US/09809665A  
; Patent No. 6790950  
; GENERAL INFORMATION:  
; APPLICANT: Lowery E., David, et al.  
; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions  
; FILE REFERENCE: 28341/00435  
; CURRENT APPLICATION NUMBER: US/09/809,665A  
; CURRENT FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 60/153,453  
; PRIOR FILING DATE: 1999-09-10  
; PRIOR APPLICATION NUMBER: 60/128,689  
; PRIOR FILING DATE: 1999-04-09  
; PRIOR APPLICATION NUMBER: 09/545,199  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 197  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 70  
; LENGTH: 3501  
; TYPE: DNA  
; ORGANISM: Pasteurella multocida  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (298)..(1905)  
; OTHER INFORMATION: yabk  
US-09-809-665A-70

Query Match 63.0%; Score 20.8; DB 4; Length 3501;  
Best Local Similarity 78.1%; Pred. No. 20;  
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGGTTTGCTTAC 33  
DB 1216 CTAACAATACGCTAGTTCCTGGTTTGCTTTC 1247

RESULT 4  
US-09-128-155-16/c  
; Sequence 16, Application US/09128155  
; Patent No. 6117654  
; GENERAL INFORMATION:  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
; FILE REFERENCE: 09404/052001  
; CURRENT APPLICATION NUMBER: US/09/128,155  
; CURRENT FILING DATE: 1998-08-03  
; EARLIER APPLICATION NUMBER: US 60/091,650  
; EARLIER FILING DATE: 1998-07-02  
; EARLIER APPLICATION NUMBER: US 60/054,646  
; EARLIER FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16

; LENGTH: 152331  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(152331)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-16

Query Match 61.8%; Score 20.4; DB 3; Length 152331;  
Best Local Similarity 80.0%; Pred. No. 66;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGGTTTGCTT 31  
DB 41497 CCACTAATGAAGTTAATTCCTGTTTGCTT 41468

RESULT 5  
US-09-128-155-17/c  
; Sequence 17, Application US/09128155  
; Patent No. 6117654  
; GENERAL INFORMATION:  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
; FILE REFERENCE: 09404/052001  
; CURRENT APPLICATION NUMBER: US/09/128,155  
; CURRENT FILING DATE: 1998-08-03  
; EARLIER APPLICATION NUMBER: US 60/091,650  
; EARLIER FILING DATE: 1998-07-02  
; EARLIER APPLICATION NUMBER: US 60/054,646  
; EARLIER FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 17  
; LENGTH: 176373  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(176373)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-17

Query Match 61.8%; Score 20.4; DB 3; Length 176373;  
Best Local Similarity 80.0%; Pred. No. 68;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGGTTTGCTT 31  
DB 11029 CCACTAATGAAGTTAATTCCTGTTTGCTT 11000

RESULT 6  
US-08-956-171E-421  
; Sequence 421, Application US/08956171E  
; Patent No. 6593114  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; Gil H. Choi  
; Patrick S. Dillon  
; Craig A. Rosen  
; Steven C. Barash  
; Michael R. Fannon  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5256  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA



COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171E  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 421:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2710 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 421:  
US-08-956-171E-421

Query Match 60.0%; Score 19.8; DB 4; Length 2710;  
Best Local Similarity 75.0%; Pred. No. 52;  
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTCGTTGCTTAC 33  
|||||  
DB 558 CCAATTAGTACGATGNTTTCGGCTTGATTAC 589

RESULT 7  
US-08-781-986A-421  
; Sequence 421, Application US/08781986A  
; Patent No. 6737248  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunesh  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5255  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 421:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2710 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-421  
Query Match 60.0%; Score 19.8; DB 4; Length 2710;  
Best Local Similarity 75.0%; Pred. No. 52;  
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 2 CCATTATTACGCTAGTTCCTCGTTGCTTAC 33  
|||||  
DB 558 CCAATTAGTACGATGNTTTCGGCTTGATTAC 589

RESULT 8  
US-09-949-016-13512/c  
; Sequence 13512, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13512  
; LENGTH: 13445  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13512

Query Match 59.4%; Score 19.6; DB 4; Length 13445;  
Best Local Similarity 84.6%; Pred. No. 89;  
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTCGTTT 27  
|||||  
DB 2022 CCGTAATCCGCTAGTTCCTGGATT 1997

RESULT 9  
US-09-949-016-107885/c  
; Sequence 107885, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 107885  
; LENGTH: 601

```
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-107885

Query Match      58.8%; Score 19.4; DB 4; Length 601;
Best Local Similarity 79.3%; Pred. No. 56;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TTTATTACGCTAGTTCCTCGTTGCTTAC 33
DB 567 TTTAAAAAAGTAGTTCCTAGTTTCTTCTAC 539

RESULT 10
US-09-949-016-14981
; Sequence 14981, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14981
; LENGTH: 115388
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14981

Query Match      58.8%; Score 19.4; DB 4; Length 115388;
Best Local Similarity 79.3%; Pred. No. 1.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTCGTTGCT 30
DB 77478 CAATCTAATGCTAGTTCCTCGTTGCT 77506

RESULT 11
US-09-949-016-14982
; Sequence 14982, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14982
; LENGTH: 127771
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14982

Query Match      58.8%; Score 19.4; DB 4; Length 127771;
Best Local Similarity 79.3%; Pred. No. 1.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTCGTTGCT 30
DB 77478 CAATCTAATGCTAGTTCCTCGTTGCT 77506

RESULT 12
US-09-949-016-14720
; Sequence 14720, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14720
; LENGTH: 390890
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(390890)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14720

Query Match      58.8%; Score 19.4; DB 4; Length 390890;
Best Local Similarity 79.3%; Pred. No. 2.1e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TTTATTACGCTAGTTCCTCGTTGCTTAC 33
DB 210696 TTTAAAAAAGTAGTTCCTAGTTTCTTCTAC 210724

RESULT 13
US-09-543-681A-2150/c
; Sequence 2150, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2150
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2150

Query Match      58.2%; Score 19.2; DB 4; Length 660;
Best Local Similarity 75.0%; Pred. No. 70;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTCGTTGCTTAC 33
DB 617 CTTTTGTTAAGGTACTTCTGATTTTTC 586
```

RESULT 14  
US-09-949-016-12421  
; Sequence 12421, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12421  
; LENGTH: 65518  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12421

Query Match 58.2%; Score 19.2; DB 4; Length 65518;  
Best Local Similarity 87.5%; Pred. No. 1.8e+02;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 ATTATTACGCTAGTCTCTGGTTT 27  
||||| ||| ||||| ||||| |||||  
DB 14264 ATTATCAGATAGTTCACGGTTT 14287

RESULT 15  
US-09-949-016-14101  
; Sequence 14101, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14101  
; LENGTH: 65518  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-14101

Query Match 58.2%; Score 19.2; DB 4; Length 65518;  
Best Local Similarity 87.5%; Pred. No. 1.8e+02;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 ATTATTACGCTAGTCTCTGGTTT 27  
||||| ||| ||||| ||||| |||||  
DB 14264 ATTATCAGATAGTTCACGGTTT 14287

RESULT 16  
US-09-949-016-16519  
; Sequence 16519, Application US/09949016

; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16519  
; LENGTH: 67386  
; TYPE: DNA  
; ORGANISM: Human  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(67386)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16519

Query Match 58.2%; Score 19.2; DB 4; Length 67386;  
Best Local Similarity 87.5%; Pred. No. 1.9e+02;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 ATTATTACGCTAGTCTCTGGTTT 27  
||||| ||| ||||| ||||| |||||  
DB 1669 ATTATTACGCTAGTCTCTGGTTT 1692

RESULT 17  
US-09-949-016-190746  
; Sequence 190746, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 190746  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-190746

Query Match 57.6%; Score 19; DB 4; Length 601;  
Best Local Similarity 81.5%; Pred. No. 84;  
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 TTATTACGCTAGTCTCTGGTTT 32  
||||| ||| ||||| ||||| |||||  
DB 142 TTATGACCTAGTCTCTGGTTT 168

RESULT 18  
US-09-949-016-17196/c  
; Sequence 17196, Application US/09949016  
; Patent No. 6812339

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17196
; LENGTH: 118143
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17196

Query Match 57.6%; Score 19; DB 4; Length 118143;
Best Local Similarity 81.5%; Pred. No. 2.5e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 TTATTACGCTAGTTCCTGGTTGCTTA 32
DB 47877 TTTGACCCCTAGTTCCTGGTTTCTTA 47851

RESULT 19
US-09-242-901-4
; Sequence 4, Application US/09242901
; Patent No. 6322783
; GENERAL INFORMATION:
; APPLICANT: TAKAHASHI, SEISHI
; TITLE OF INVENTION: NOVEL BACTERIOPHAGES, METHOD FOR SCREENING SAME AND
; DETECTION OF BACTERIOPHAGES IN BACTERIAL COMPOSITIONS USING SAME, AND DETECTION
; OF BACTERIOPHAGES IN BACTERIAL COMPOSITIONS USING SAME
; FILE REFERENCE: 3000-0001
; CURRENT APPLICATION NUMBER: US/09/242,901
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: PCT/JP97/02957
; EARLIER FILING DATE: 1997-08-26
; EARLIER APPLICATION NUMBER: 8/261132
; EARLIER FILING DATE: 1996-08-26
; EARLIER APPLICATION NUMBER: 9/135716
; EARLIER FILING DATE: 1997-04-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Bacteriophage
; FEATURE:
; OTHER INFORMATION: "n" at various positions throughout the sequence may be
; OTHER INFORMATION: a, t, g, c other or unknown
US-09-242-901-4

Query Match 57.0%; Score 18.8; DB 3; Length 599;
Best Local Similarity 76.7%; Pred. No. 1e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGGTTGCTT 31
DB 165 CCATTAAATCTAGTTCCTGGTTTCA 194

RESULT 20
US-09-072-917A-8/c
; Sequence 8, Application US/09072917A
; Patent No. 6288302
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```
; APPLICANT: Yu, Su-May
; APPLICANT: Liu, Li-Fei
; APPLICANT: Chan, Ming-Tsair
; TITLE OF INVENTION: Application of Alpha-Amylase Gene
; TITLE OF INVENTION: Promoter and Signal Sequence in the Production of
; Patent No. 6288302
; TITLE OF INVENTION: Recombinant Proteins in Transgenic Plants and Transgenic
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/072,917A
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/509,962
; FILING DATE: 01-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28123/34257
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2938 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Rice (Oryzae sativa)
; STRAIN:
; IMMEDIATE SOURCE:
; LIBRARY: (EMBL) genomic
; CLONE: alpha-Amy3-C
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1081..1170, 1295..2230, 2319..2600)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: join(1156..1170, 1295..2230, 2319..2600)
US-09-072-917A-8

Query Match 57.0%; Score 18.8; DB 3; Length 2938;
Best Local Similarity 76.7%; Pred. No. 1.4e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 ATTTATTACGCTAGTTCCTGGTTGCTTAC 33
DB 2260 ATATAACTATAATGATGTTTCTTAC 2231

RESULT 21
US-09-949-016-17486
; Sequence 17486, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17486  
; LENGTH: 56976  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(56976)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-17486

Query Match 57.0%; Score 18.8; DB 4; Length 56976;  
Best Local Similarity 76.7%; Pred. No. 2.7e+02;  
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 ATTATTACGCTAGTTCCTGTTGCTTAC 33  
|||||  
Db 53416 ATTTTGTCTTTGTTCTGTTGCTTGC 53445

RESULT 22  
US-09-949-016-12033  
; Sequence 12033, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12033  
; LENGTH: 101674  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(101674)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12033

Query Match 57.0%; Score 18.8; DB 4; Length 101674;  
Best Local Similarity 76.7%; Pred. No. 3e+02;  
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 ATTATTACGCTAGTTCCTGTTGCTTAC 33  
|||||  
Db 53415 ATTTTGTCTTTGTTCTGTTGCTTGC 53444

RESULT 23  
US-09-949-016-12656  
; Sequence 12656, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12656  
; LENGTH: 253345  
; TYPE: DNA  
; ORGANISM: Human  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12656

Query Match 57.0%; Score 18.8; DB 4; Length 253345;  
Best Local Similarity 76.7%; Pred. No. 3.6e+02;  
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGTTGCTTA 32  
|||||  
Db 59674 CAGTTATTAGGCTACTGCTTGATTGATTA 59703

RESULT 24  
US-09-949-016-13639  
; Sequence 13639, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13639  
; LENGTH: 253364  
; TYPE: DNA  
; ORGANISM: Human  
; OTHER INFORMATION:  
US-09-949-016-13639

Query Match 57.0%; Score 18.8; DB 4; Length 253364;  
Best Local Similarity 76.7%; Pred. No. 3.6e+02;  
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGTTGCTTA 32  
|||||  
Db 59674 CAGTTATTAGGCTACTGCTTGATTGATTA 59703

RESULT 25  
US-09-949-016-12386  
; Sequence 12386, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12386  
; LENGTH: 263693  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12386

Query Match 57.0%; Score 18.8; DB 4; Length 263693;  
Best Local Similarity 76.7%; Pred. No. 3.6e+02;  
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 CATTATTACGCTAGTTCCTGCTTGCTTA 32  
Db 26393 CATTATTATAGTAGTTCATAATTAGCTTA 26422

RESULT 26  
US-09-949-016-16915  
; Sequence 16915, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16915  
; LENGTH: 263694  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16915

Query Match 57.0%; Score 18.8; DB 4; Length 263694;  
Best Local Similarity 76.7%; Pred. No. 3.6e+02;  
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 CATTATTACGCTAGTTCCTGCTTGCTTA 32  
Db 26393 CATTATTATAGTAGTTCATAATTAGCTTA 26422

RESULT 27  
US-08-400-256-37/c  
; Sequence 37, Application US/08400256  
; Patent No. 5750497  
; GENERAL INFORMATION:  
; APPLICANT: Havelund, Svend  
; APPLICANT: Halstrom, John  
; APPLICANT: Jonassen, Ib  
; APPLICANT: Andersen, Asger Sloth  
; APPLICANT: Markussen, Jan  
; TITLE OF INVENTION: ACYLATED INSULIN  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5750497o No. 5750497d/disk of No. 5750497th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor

; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/400,256  
; FILING DATE: 03-MAR-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J.  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 3985.220-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 409 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-400-256-37

Query Match 56.4%; Score 18.6; DB 1; Length 409;  
Best Local Similarity 84.0%; Pred. No. 1.2e+02;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 TATTACGCTAGTTCCTGCTTGCTT 31  
Db 119 TCTTAGGCTAGTTCCTGTTTGGTT 95

RESULT 28  
US-08-975-365-37/c  
; Sequence 37, Application US/08975365  
; Patent No. 6011007  
; GENERAL INFORMATION:  
; APPLICANT: Havelund, Svend  
; APPLICANT: Halstrom, John  
; APPLICANT: Jonassen, Ib  
; APPLICANT: Andersen, Asger Sloth  
; APPLICANT: Markussen, Jan  
; TITLE OF INVENTION: ACYLATED INSULIN  
; NUMBER OF SEQUENCES: 49  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6011007o No. 6011007d/disk of No. 6011007th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10174-6401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/975,365  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/400,256  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J.  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 3985.220-US

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 409 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-975-365-37

Query Match          56.4%; Score 18.6; DB 3; Length 409;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 TATTACGCTAGTCTCTGGTTTGCTT 31
   |||||
Db 119 TCTTAGGCTAGTCTCTGGTTTGCTT 95

RESULT 29
US-08-400-256-16/c
; Sequence 16, Application US/08400256
; Patent No. 5750497
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5750497o No. 5750497disk of No. 5750497th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,256
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-400-256-22

Query Match          56.4%; Score 18.6; DB 1; Length 415;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 TATTACGCTAGTCTCTGGTTTGCTT 31
   |||||
Db 119 TCTTAGGCTAGTCTCTGGTTTGCTT 95

RESULT 31
US-08-400-256-25/c
; Sequence 25, Application US/08400256
; Patent No. 5750497
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5750497o No. 5750497disk of No. 5750497th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,256
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-400-256-16

Query Match          56.4%; Score 18.6; DB 1; Length 415;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 TATTACGCTAGTCTCTGGTTTGCTT 31
   |||||
Db 119 TCTTAGGCTAGTCTCTGGTTTGCTT 95
```

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,256
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-400-256-25

Query Match 56.4%; Score 18.6; DB 1; Length 415;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 TATTACGCTAGTTCCTGTTTGCTT 31
Db 119 TCTTAGGCTAGTTCCTGTTTGCTT 95

RESULT 32
US-08-400-256-28/c
; Sequence 28, Application US/08400256
; Patent No. 5750497
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5750497o No. 5750497disk of No. 5750497th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,256
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-400-256-31

Query Match 56.4%; Score 18.6; DB 1; Length 415;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 TATTACGCTAGTTCCTGTTTGCTT 31
Db 119 TCTTAGGCTAGTTCCTGTTTGCTT 95

RESULT 34
US-08-975-365-16/c
; Sequence 16, Application US/08975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
```



```

; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-975-365-16

Query Match 56.4%; Score 18.6; DB 3; Length 415;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 TATTAGGCTAGTTCCTGGTTTGCTT 31
Db 119 TCTTAGGCTAGTTCCTGGTTTGCTT 95

RESULT 35
US-08-975-365-22/c
; Sequence 22, Application US/08975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-975-365-22

Query Match 56.4%; Score 18.6; DB 3; Length 415;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 TATTAGGCTAGTTCCTGGTTTGCTT 31
Db 119 TCTTAGGCTAGTTCCTGGTTTGCTT 95

RESULT 36
US-08-975-365-25/c
; Sequence 25, Application US/08975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-975-365-22
```

```
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-975-365-25

Query Match          56.4%; Score 18.6; DB 3; Length 415;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 TATTACGCTAGTTCCTGGTTTGCTT 31
   |||||
Db 119 TCTTAGGCTAGTTCCTGTTTGTT 95

RESULT 37
US-08-975-365-28/c
; Sequence 28, Application US/08975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-975-365-28

Query Match          56.4%; Score 18.6; DB 3; Length 415;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 TATTACGCTAGTTCCTGGTTTGCTT 31
   |||||
Db 119 TCTTAGGCTAGTTCCTGTTTGTT 95

RESULT 38
US-08-975-365-28/c
; Sequence 28, Application US/08975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-975-365-28

Query Match          56.4%; Score 18.6; DB 3; Length 415;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 TATTACGCTAGTTCCTGGTTTGCTT 31
   |||||
Db 119 TCTTAGGCTAGTTCCTGTTTGTT 95

RESULT 39
US-09-949-016-206088/c
; Sequence 206088, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 206088
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-206088
```

```
Query Match          56.4%; Score 18.6; DB 4; Length 601;
Best Local Similarity 72.7%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCCATTATTACGCTAGTTCCTGGTTGCTTAC 33
   ||||| || ||||| || ||||| |||||
Db 234 CTCATTCTTTGCTATTACTAGTTATTAT 202
```

```
RESULT 40
US-09-949-016-206125/c
; Sequence 206125, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 206125
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-206125
```

```
Query Match          56.4%; Score 18.6; DB 4; Length 601;
Best Local Similarity 72.7%; Pred. No. 1.3e+02;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 CCCATTATTACGCTAGTTCCTGGTTGCTTAC 33
   ||||| || ||||| || ||||| |||||
Db 234 CTCATTCTTTGCTATTACTAGTTATTAT 202
```

Search completed: June 4, 2005, 15:59:14  
Job time : 69 secs

**This Page Blank (Usnta)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 14:31:52 ; Search time 261 Seconds  
(without alignments)  
777.260 Million cell updates/sec

Title: US-10-089-452-26

Perfect score: 33

Sequence: 1 cccattattacgtactgcttcgttgcttac 33

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21.4	64.8	161671	15	US-10-017-117-1
2	20.8	63.0	757	13	Sequence 1, Appli
3	20.8	63.0	757	13	Sequence 164135,
4	20.8	63.0	3501	11	US-10-027-632-164135
5	20.8	63.0	3501	11	Sequence 70, Appl
6	20.4	61.8	617	17	US-10-854-299-70
7	20.4	61.8	1240	17	Sequence 70, Appl
8	20.4	61.8	1240	17	Sequence 2746, Ap
9	20.4	61.8	1488	17	US-10-374-780A-686
10	20.4	61.8	1488	17	Sequence 686, App
11	20.4	61.8	152331	13	US-10-412-699B-1139
12	20.4	61.8	152331	13	Sequence 1139, Ap
13	20.4	61.8	176373	13	US-10-282-122A-7061
14	20.4	61.8	176373	13	Sequence 7061, Ap
15	20.4	61.8	176373	13	Sequence 15, Appl
16	20.4	61.8	176373	13	Sequence 17, Appl

12	20.2	61.2	1006	17	US-10-371-701-9	Sequence 9, Appli
13	20.2	61.2	102790	18	US-10-367-094-163	Sequence 163, App
14	20	60.6	228	18	US-10-425-115-19223	Sequence 19223, A
15	20	60.6	1203	18	US-10-483-241-11	Sequence 11, Appl
16	20	60.6	1203	18	US-10-483-241-13	Sequence 13, Appl
17	20	60.6	1637	18	US-10-483-241-14	Sequence 14, Appl
18	20	60.6	2464	18	US-10-483-241-9	Sequence 9, Appli
19	20	60.6	3761	18	US-10-483-241-6	Sequence 6, Appli
20	20	60.6	10735	18	US-10-483-241-5	Sequence 5, Appli
21	19.8	60.0	483	13	US-10-027-632-88930	Sequence 88930, A
22	19.8	60.0	483	13	US-10-027-632-303567	Sequence 303567,
23	19.8	60.0	483	17	US-10-027-632-88930	Sequence 88930, A
24	19.8	60.0	483	17	US-10-027-632-303567	Sequence 303567,
25	19.8	60.0	2710	8	US-08-781-986A-421	Sequence 421, App
26	19.8	60.0	2710	17	US-10-329-624-421	Sequence 421, App
27	19.8	60.0	31814	9	US-09-817-182-3	Sequence 3, Appli
28	19.8	60.0	31814	18	US-10-473-349-3	Sequence 3, Appli
29	19.8	60.0	39121	13	US-10-087-192-481	Sequence 481, App
30	19.8	60.0	653122	13	US-10-087-192-226	Sequence 226, App
31	19.6	59.4	310	18	US-10-723-860-1322	Sequence 1322, Ap
32	19.4	58.8	903	17	US-10-369-493-26447	Sequence 26447, A
33	19.4	58.8	518360	18	US-10-367-094-125	Sequence 125, App
34	19.4	58.8	653458	19	US-10-461-862-4	Sequence 4, Appli
35	19.2	58.2	354	14	US-10-056-052-15	Sequence 15, Appl
36	19.2	58.2	411	18	US-10-425-115-23753	Sequence 23753, A
37	19.2	58.2	540	16	US-10-161-051-101	Sequence 101, App
38	19.2	58.2	909	17	US-10-282-122A-34217	Sequence 34217, A
39	19.2	58.2	1839	17	US-10-282-122A-36283	Sequence 36283, A
40	19.2	58.2	2218	17	US-10-425-114-33630	Sequence 33630, A
41	19.2	58.2	493631	13	US-10-087-192-205	Sequence 205, App
42	19	57.6	354	14	US-10-056-052-7	Sequence 7, Appli
43	19	57.6	717	9	US-09-770-149-173	Sequence 173, App
44	19	57.6	100554	13	US-10-087-192-1165	Sequence 1165, Ap
45	18.8	57.0	465	18	US-10-437-963-81374	Sequence 81374, A

ALIGNMENTS

RESULT 1

US-10-017-117-1  
; Sequence 1, Application US/10017117  
; Publication No. US20030124535A1  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Jeanette  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE  
; FILE REFERENCE: MMI-008  
; CURRENT APPLICATION NUMBER: US/10/017,117  
; CURRENT FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: NO. US20030124535A1 yet assigned  
; PRIOR FILING DATE: 2001-12-05  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 161671  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(161671)  
; OTHER INFORMATION: n = a,t,c or g  
US-10-017-117-1

Query Match	64.8%	Score 21.4;	DB 15;	Length 161671;
Best Local Similarity	80.6%	Pred. No. 1.6e+02;		
Matches 25;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	1	CCCATTTATTACGCTAGTTCCTGTTGCTT	31	
DB	125929	CCCATTTAAAACTCTATTTCATTGTTGCTT	125959	
RESULT 2				

```
US-10-027-632-164135
; Sequence 164135, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164135
; LENGTH: 757
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-164135

Query Match          63.0%; Score 20.8; DB 13; Length 757;
Best Local Similarity 78.1%; Pred. No. 97;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCCATTATTACGCTAGTTCTCGGTTTGCTTA 32
Db 135 CGCATTATGACTCTATTTCTGTTTACCTTA 166

RESULT 3
US-10-027-632-164135
; Sequence 164135, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164135
; LENGTH: 757
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-164135
```

```
Query Match          63.0%; Score 20.8; DB 17; Length 757;
Best Local Similarity 78.1%; Pred. No. 97;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCCATTATTACGCTAGTTCTCGGTTTGCTTA 32
Db 135 CGCATTATGACTCTATTTCTGTTTACCTTA 166

RESULT 4
US-09-809-665A-70
; Sequence 70, Application US/09809665A
; Publication No. US20040110268A1
; GENERAL INFORMATION:
; APPLICANT: Lowery E., David, et al.
; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
; FILE REFERENCE: 28341/00435
; CURRENT APPLICATION NUMBER: US/09/809,665A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/153,453
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/128,689
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/545,199
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; LENGTH: 3501
; TYPE: DNA
; ORGANISM: Pasteurella multocida
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (298)..(1905)
; OTHER INFORMATION: yabk
US-09-809-665A-70

Query Match          63.0%; Score 20.8; DB 11; Length 3501;
Best Local Similarity 78.1%; Pred. No. 1.3e+02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCTCGGTTTGCTTAC 33
Db 1216 CTACTCAATACGCTAGTTCTGCTTTGCTTTC 1247

RESULT 5
US-10-854-299-70
; Sequence 70, Application US/10854299
; Publication No. US20050003512A1
; GENERAL INFORMATION:
; APPLICANT: Lowery E., David, et al.
; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
; FILE REFERENCE: 28341/00435
; CURRENT APPLICATION NUMBER: US/10/854,299
; CURRENT FILING DATE: 2004-05-26
; PRIOR APPLICATION NUMBER: 60/153,453
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/128,689
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/545,199
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; LENGTH: 3501
; TYPE: DNA
; ORGANISM: Pasteurella multocida
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (298)..(1905)
; OTHER INFORMATION: yabk
```

OTHER INFORMATION: yabk

US-10-854-299-70

Query Match 63.0%; Score 20.8; DB 18; Length 3501;  
Best Local Similarity 78.1%; Pred. No. 1.3e+02;  
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGCTTGCCTTAC 33  
Db 1216 CTACTCAATACGCTAGTTCCTGCTTGCCTTTC 1247

## RESULT 6

US-10-282-122A-2746  
Sequence 2746, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2746

LENGTH: 617

TYPE: DNA

ORGANISM: Escherichia coli

US-10-282-122A-2746

Query Match 61.8%; Score 20.4; DB 17; Length 617;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 ATTATTACGCTAGTTCCTGCTTGCCTTAC 33

Db 107 AGTCGTTACGCTAGTTCCTGCTTGCCTTAC 136

## RESULT 7

US-10-374-780A-686/c

Sequence 686, Application US/10374780A

Publication No. US20040019927A1

GENERAL INFORMATION:

APPLICANT: Sherman, Bradley K

APPLICANT: Riechmann, Jose Luis

APPLICANT: Jiang, Cai-Zhong

APPLICANT: Heard, Jacqueline E

APPLICANT: Haake, Volker

APPLICANT: Creelman, Robert A

APPLICANT: Ratcliffe, Oliver

APPLICANT: Adam, Luc J

APPLICANT: Reuber, T. Lynne

APPLICANT: Keddle, James

APPLICANT: Broun, Pierre E

APPLICANT: Filgrim, Marsha L

APPLICANT: Dubell III, Arnold T

APPLICANT: Pineda, Omaira

APPLICANT: Yu, Guo-Liang

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

FILE REFERENCE: MBI-0047 CIP

CURRENT APPLICATION NUMBER: US/10/374,780A

CURRENT FILING DATE: 2003-02-25

PRIOR APPLICATION NUMBER: 09/837,944

PRIOR FILING DATE: 2001-04-18

PRIOR APPLICATION NUMBER: 60/310,847

PRIOR FILING DATE: 2001-08-09

PRIOR APPLICATION NUMBER: 09/934,455

PRIOR FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: 60/336,049

PRIOR FILING DATE: 2001-11-19

PRIOR APPLICATION NUMBER: 60/338,692

PRIOR FILING DATE: 2001-12-11

PRIOR APPLICATION NUMBER: 10/171,468

PRIOR FILING DATE: 2002-06-14

PRIOR APPLICATION NUMBER: 10/225,066

PRIOR FILING DATE: 2002-08-09

PRIOR APPLICATION NUMBER: 10/225,067

PRIOR FILING DATE: 2002-08-09

PRIOR APPLICATION NUMBER: 10/225,068

PRIOR FILING DATE: 2002-08-09

NUMBER OF SEQ ID NOS: 2906

SOFTWARE: PatentIn version 3.2

SEQ ID NO 686

LENGTH: 1240

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G409

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (680)..(680)

OTHER INFORMATION: n is a, c, g, or t

US-10-374-780A-686

Query Match 61.8%; Score 20.4; DB 17; Length 1240;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCCATTATTACGCTAGTTCCTGCTTGTCT 30

Db 1101 CCCAGTTTAACTCTAGTTCCTAGTTTCCT 1072

## RESULT 8

US-10-412-699B-1139/c

Sequence 1139, Application US/10412699B

Publication No. US20040045049A1

GENERAL INFORMATION:

APPLICANT: Mendel Biotechnology, Inc.

APPLICANT: Zhang, James

APPLICANT: Fromm, Michael E.

APPLICANT: Heard, Jacqueline E.

APPLICANT: Riechmann, Jose Luis

APPLICANT: Adam, Luc J.  
APPLICANT: Broun, Pierre E.  
APPLICANT: Pineda, Omaira  
APPLICANT: Reuber, T. Lynne  
APPLICANT: Keddle, James S.  
APPLICANT: Yu, Guo-Liang  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Samaha, Raymond R.  
APPLICANT: Pilgrim, Marsha L.  
APPLICANT: Creelman, Robert A.  
APPLICANT: Dubell, Arnold N.  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Katimoto, Roderick  
APPLICANT: Sherman, Bradley K.  
TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants  
FILE REFERENCE: MBI-0048CIP  
CURRENT APPLICATION NUMBER: US/10/412,699B  
CURRENT FILING DATE: 2003-04-10  
PRIOR APPLICATION NUMBER: 09/394,519  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: 09/489,376  
PRIOR FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: 09/506,720  
PRIOR FILING DATE: 2000-02-17  
PRIOR APPLICATION NUMBER: 09/533,030  
PRIOR FILING DATE: 2000-03-22  
PRIOR APPLICATION NUMBER: 09/533,392  
PRIOR FILING DATE: 2000-03-22  
PRIOR APPLICATION NUMBER: 09/533,029  
PRIOR FILING DATE: 2000-03-22  
PRIOR APPLICATION NUMBER: 09/532,591  
PRIOR FILING DATE: 2000-03-22  
PRIOR APPLICATION NUMBER: 09/533,648  
PRIOR FILING DATE: 2000-03-22  
PRIOR APPLICATION NUMBER: 09/713,994  
PRIOR FILING DATE: 2000-11-16  
PRIOR APPLICATION NUMBER: 09/819,142  
PRIOR FILING DATE: 2001-03-27  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 2011  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 1139  
LENGTH: 1240  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (680)..(680)  
OTHER INFORMATION: n is a, c, g, or t  
US-10-412-699B-1139

Query Match 61.8%; Score 20.4; DB 17; Length 1240;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCCATTATTACGCTAGTTCCTGGTTTGGT 30  
|||||  
DB 1101 CCCAGTTTAACTCTAGTTCCTAGTTTCT 1072

RESULT 9  
US-10-282-122A-7061/c  
Sequence 7061, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Lianguo  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John

APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7061  
LENGTH: 1488  
TYPE: DNA  
ORGANISM: Escherichia coli  
US-10-282-122A-7061

Query Match 61.8%; Score 20.4; DB 17; Length 1488;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 ATTATTACGCTAGTTCCTGGTTTGGCTTAC 33  
|||||  
DB 1447 AGTCGTACGCTCGTTACAGGTTTGGTTAC 1418

RESULT 10  
US-10-095-407-16/c  
Sequence 16, Application US/10095407  
Publication No. US20020164330A1  
GENERAL INFORMATION:  
APPLICANT: Pan, Yang  
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
FILE REFERENCE: 09404/052001  
CURRENT APPLICATION NUMBER: US/10/095,407  
CURRENT FILING DATE: 2002-03-11  
PRIOR APPLICATION NUMBER: US 60/091,650  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: US 60/054,646  
PRIOR FILING DATE: 1997-08-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 16  
LENGTH: 152331  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)-(152331)  
OTHER INFORMATION: n = A,T,C or G  
US-10-095-407-16



Query Match 61.8%; Score 20.4; DB 13; Length 152331;  
Best Local Similarity 80.0%; Pred. No. 4.1e+02;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGTTGCTT 31  
|||||  
DB 41497 CCATCTATGAAGTTAAATTCCTGTTTGCTT 41468  
|||||

RESULT 11  
US-10-095-407-17/c  
; Sequence 17, Application US/10095407  
; Publication No. US20020164330A1  
; GENERAL INFORMATION:  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
; FILE REFERENCE: 09404/052001  
; CURRENT APPLICATION NUMBER: US/10/095,407  
; CURRENT FILING DATE: 2002-03-11  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: US 60/091,650  
; PRIOR FILING DATE: 1997-08-04  
; PRIOR APPLICATION NUMBER: US 60/054,646  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 17  
; LENGTH: 176373  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(176373)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-095-407-17

Query Match 61.8%; Score 20.4; DB 13; Length 176373;  
Best Local Similarity 80.0%; Pred. No. 4.2e+02;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGTTGCTT 31  
|||||  
DB 11029 CCATCTATGAAGTTAAATTCCTGTTTGCTT 11000  
|||||

RESULT 12  
US-10-371-701-9  
; Sequence 9, Application US/10371701  
; Publication No. US20040091985A1  
; GENERAL INFORMATION:  
; APPLICANT: METCALF, WILLIAM W.  
; APPLICANT: VAN DER DONK, WILFRED A.  
; APPLICANT: VRTIS, JENNIFER M.  
; APPLICANT: WHITE, ANDREA K.  
; APPLICANT: GARCIA COSTAS, AMAYA M.  
; TITLE OF INVENTION: NAD PHOSPHITE OXIDOREDUCTASE, A NOVEL CATALYST FROM  
; FILE REFERENCE: 21419/94021  
; CURRENT APPLICATION NUMBER: US/10/371,701  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: 60/359,091  
; PRIOR FILING DATE: 2002-02-22  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 1006  
; TYPE: DNA  
; ORGANISM: Alcaligenes faecalis  
US-10-371-701-9

Query Match 61.2%; Score 20.2; DB 17; Length 1006;  
Best Local Similarity 88.0%; Pred. No. 1.8e+02;  
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 TTACGCTAGTTCCTGTTGCTTAC 33  
|||||  
DB 278 TTACGATTGTTCTGATTGCTTAC 302  
|||||

RESULT 13  
US-10-367-094-163/c  
; Sequence 163, Application US/10367094  
; Publication No. US20040170982A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer  
; FILE REFERENCE: 529452001500  
; CURRENT APPLICATION NUMBER: US/10/367,094  
; CURRENT FILING DATE: 2003-02-14  
; NUMBER OF SEQ ID NOS: 203  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 163  
; LENGTH: 102790  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(102790)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-367-094-163

Query Match 61.2%; Score 20.2; DB 18; Length 102790;  
Best Local Similarity 75.8%; Pred. No. 4.6e+02;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCCATTATTACGCTAGTTCCTGTTGCTTAC 33  
|||||  
DB 100253 CCCCTTATTACCTTTGTACCTTTTGTCTGC 100221  
|||||

RESULT 14  
US-10-425-115-19223  
; Sequence 19223, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 19223  
; LENGTH: 228  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_117533C.1  
US-10-425-115-19223

Query Match 60.6%; Score 20; DB 18; Length 228;  
Best Local Similarity 82.1%; Pred. No. 1.7e+02;  
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 TTATTACGCTAGTTCCTGTTGCTTAC 33  
|||||  
DB 113 TTATTTCGTTGTTCTGCTTTGCTTAC 140  
|||||

RESULT 15  
US-10-483-241-11  
; Sequence 11, Application US/10483241

; Publication No. US20040259251A1  
; GENERAL INFORMATION:  
; APPLICANT: Broggin, Massimo  
; APPLICANT: D'Incalci, Maurizio  
; TITLE OF INVENTION: Oncosuppressive Gene  
; FILE REFERENCE: 2965-187  
; CURRENT APPLICATION NUMBER: US/10/483,241  
; PRIOR FILING DATE: 2004-01-09  
; PRIOR APPLICATION NUMBER: PCT/EP02/07625  
; PRIOR FILING DATE: 2002-07-09  
; PRIOR APPLICATION NUMBER: MI 2001A001465  
; PRIOR FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 11  
; LENGTH: 1203  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-483-241-11

Query Match 60.6%; Score 20; DB 18; Length 1203;  
Best Local Similarity 82.1%; Pred. No. 2.3e+02;  
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 TTATTACGCTAGTTCCTGTTGCTTAC 33  
DB 779 TTACTGCTAGTTCTGTTGTTTAC 806

RESULT 16  
US-10-483-241-13  
; Sequence 13, Application US/10483241  
; Publication No. US20040259251A1  
; GENERAL INFORMATION:  
; APPLICANT: Broggin, Massimo  
; APPLICANT: D'Incalci, Maurizio  
; TITLE OF INVENTION: Oncosuppressive Gene  
; FILE REFERENCE: 2965-187  
; CURRENT APPLICATION NUMBER: US/10/483,241  
; PRIOR FILING DATE: 2004-01-09  
; PRIOR APPLICATION NUMBER: PCT/EP02/07625  
; PRIOR FILING DATE: 2002-07-09  
; PRIOR APPLICATION NUMBER: MI 2001A001465  
; PRIOR FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 13  
; LENGTH: 1203  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-483-241-13

Query Match 60.6%; Score 20; DB 18; Length 1203;  
Best Local Similarity 82.1%; Pred. No. 2.3e+02;  
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 TTATTACGCTAGTTCCTGTTGCTTAC 33  
DB 779 TTACTGCTAGTTCTGTTGTTTAC 806

RESULT 17  
US-10-483-241-14  
; Sequence 14, Application US/10483241  
; Publication No. US20040259251A1  
; GENERAL INFORMATION:  
; APPLICANT: Broggin, Massimo  
; APPLICANT: D'Incalci, Maurizio  
; TITLE OF INVENTION: Oncosuppressive Gene  
; FILE REFERENCE: 2965-187  
; CURRENT APPLICATION NUMBER: US/10/483,241  
; PRIOR FILING DATE: 2004-01-09  
; PRIOR APPLICATION NUMBER: PCT/EP02/07625

; PRIOR FILING DATE: 2002-07-09  
; PRIOR APPLICATION NUMBER: MI 2001A001465  
; PRIOR FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 14  
; LENGTH: 1637  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-483-241-14

Query Match 60.6%; Score 20; DB 18; Length 1637;  
Best Local Similarity 82.1%; Pred. No. 2.5e+02;  
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 TTATTACGCTAGTTCCTGTTGCTTAC 33  
DB 779 TTACTGCTAGTTCTGTTGTTTAC 806

RESULT 18  
US-10-483-241-9  
; Sequence 9, Application US/10483241  
; Publication No. US20040259251A1  
; GENERAL INFORMATION:  
; APPLICANT: Broggin, Massimo  
; APPLICANT: D'Incalci, Maurizio  
; TITLE OF INVENTION: Oncosuppressive Gene  
; FILE REFERENCE: 2965-187  
; CURRENT APPLICATION NUMBER: US/10/483,241  
; PRIOR FILING DATE: 2004-01-09  
; PRIOR APPLICATION NUMBER: PCT/EP02/07625  
; PRIOR FILING DATE: 2002-07-09  
; PRIOR APPLICATION NUMBER: MI 2001A001465  
; PRIOR FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 9  
; LENGTH: 2464  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-483-241-9

Query Match 60.6%; Score 20; DB 18; Length 2464;  
Best Local Similarity 82.1%; Pred. No. 2.7e+02;  
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 TTATTACGCTAGTTCCTGTTGCTTAC 33  
DB 779 TTACTGCTAGTTCTGTTGTTTAC 806

RESULT 19  
US-10-483-241-6  
; Sequence 6, Application US/10483241  
; Publication No. US20040259251A1  
; GENERAL INFORMATION:  
; APPLICANT: Broggin, Massimo  
; APPLICANT: D'Incalci, Maurizio  
; TITLE OF INVENTION: Oncosuppressive Gene  
; FILE REFERENCE: 2965-187  
; CURRENT APPLICATION NUMBER: US/10/483,241  
; PRIOR FILING DATE: 2004-01-09  
; PRIOR APPLICATION NUMBER: PCT/EP02/07625  
; PRIOR FILING DATE: 2002-07-09  
; PRIOR APPLICATION NUMBER: MI 2001A001465  
; PRIOR FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 6  
; LENGTH: 3761  
; TYPE: DNA  
; ORGANISM: Homo sapiens

## US-10-483-241-6

Query Match 60.6%; Score 20; DB 18; Length 3761;  
Best Local Similarity 82.1%; Pred. No. 2.9e+02;  
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 TTATTACGCTAGTTCCTGTTGCTTAC 33  
|||||  
Db 1917 TTATACGCTAGTTCCTGTTGCTTAC 1944  
|||||

## RESULT 20

US-10-483-241-5  
; Sequence 5, Application US/10483241  
; Publication No. US20040259251A1  
; GENERAL INFORMATION:  
; APPLICANT: Broggini, Massimo  
; TITLE OF INVENTION: Oncosuppressive Gene  
; FILE REFERENCE: 2965-187  
; CURRENT APPLICATION NUMBER: US/10/483,241  
; CURRENT FILING DATE: 2004-01-09  
; PRIOR APPLICATION NUMBER: PCT/EP02/07625  
; PRIOR FILING DATE: 2002-07-09  
; PRIOR APPLICATION NUMBER: MI 2001A001465  
; PRIOR FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5  
; LENGTH: 10735  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-483-241-5

Query Match 60.6%; Score 20; DB 18; Length 10735;  
Best Local Similarity 82.1%; Pred. No. 3.6e+02;  
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 TTATTACGCTAGTTCCTGTTGCTTAC 33  
|||||  
Db 8891 TTATACGCTAGTTCCTGTTGCTTAC 8918  
|||||

## RESULT 21

US-10-027-632-88930  
; Sequence 88930, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 88930  
; LENGTH: 483  
; TYPE: DNA

; ORGANISM: Human  
US-10-027-632-88930

Query Match 60.0%; Score 19.8; DB 13; Length 483;  
Best Local Similarity 77.4%; Pred. No. 2.4e+02;  
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGTTGCTTAC 33  
|||||  
Db 290 CTTTCTAACTAGTTCCTGTTTCTTAC 320  
|||||

## RESULT 22

US-10-027-632-303567  
; Sequence 303567, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 303567  
; LENGTH: 483  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-303567

Query Match 60.0%; Score 19.8; DB 13; Length 483;  
Best Local Similarity 77.4%; Pred. No. 2.4e+02;  
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGTTGCTTAC 33  
|||||  
Db 290 CTTTCTAACTAGTTCCTGTTTCTTAC 320  
|||||

## RESULT 23

US-10-027-632-88930  
; Sequence 88930, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24

;; PRIOR APPLICATION NUMBER: US 60/167,363  
;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358  
;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002  
;; PRIOR FILING DATE: 1999-08-09  
;; NUMBER OF SEQ ID NOS: 325720  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 88930  
;; LENGTH: 483  
;; TYPE: DNA  
;; ORGANISM: Human  
US-10-027-632-88930

Query Match 60.0%; Score 19.8; DB 17; Length 483;  
Best Local Similarity 77.4%; Pred. No. 2.4e+02;  
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGTTGCTTAC 33  
Db 290 CTTTCTTAACCTAGTTCCTGTTTCTTAC 320

RESULT 24  
US-10-027-632-303567  
;; Sequence 303567, Application US/10027632  
;; Publication No. US20030204075A9  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
;; TITLE OF INVENTION: Polymorphisms in the Human Genome  
;; FILE REFERENCE: 108827.129  
;; CURRENT APPLICATION NUMBER: US/10/027,632  
;; CURRENT FILING DATE: 2002-04-30  
;; PRIOR APPLICATION NUMBER: US 60/218,006  
;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/198,676  
;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483  
;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363  
;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358  
;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002  
;; PRIOR FILING DATE: 1999-08-09  
;; NUMBER OF SEQ ID NOS: 325720  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 303567  
;; LENGTH: 483  
;; TYPE: DNA  
;; ORGANISM: Human  
US-10-027-632-303567

Query Match 60.0%; Score 19.8; DB 17; Length 483;  
Best Local Similarity 77.4%; Pred. No. 2.4e+02;  
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGTTGCTTAC 33  
Db 290 CTTTCTTAACCTAGTTCCTGTTTCTTAC 320

RESULT 25  
US-08-781-986A-421  
;; Sequence 421, Application US/08781986A  
;; Publication No. US20030054436A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Charles Kunsch  
;; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
;; NUMBER OF SEQUENCES: 5255

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Human Genome Sciences, Inc.  
;; STREET: 9410 Key West Avenue  
;; CITY: Rockville  
;; STATE: Maryland  
;; COUNTRY: USA  
;; ZIP: 20850  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
;; COMPUTER: HP Vectra 486/33  
;; OPERATING SYSTEM: MSDOS version 6.2  
;; SOFTWARE: ASCII Text  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/781,986A  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Benson, Bob  
;; REGISTRATION NUMBER: 30,446  
;; REFERENCE/DOCKET NUMBER: PB248PP  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (301) 309-8504  
;; TELEFAX: (301) 309-8512  
;; INFORMATION FOR SEQ ID NO: 421:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2710 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
US-08-781-986A-421  
Query Match 60.0%; Score 19.8; DB 8; Length 2710;  
Best Local Similarity 75.0%; Pred. No. 3.3e+02;  
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 2 CCATTATTACGCTAGTTCCTGTTGCTTAC 33  
Db 558 CCATTATTACGCTAGTTCCTGTTGCTTAC 589  
RESULT 26  
US-10-329-624-421  
;; Sequence 421, Application US/10329624  
;; Publication No. US20040043037A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Charles Kunsch  
;; Gil H. Choi  
;; Patrick S. Dillon  
;; Craig A. Rosen  
;; Steven C. Barash  
;; Michael R. Fannon  
;; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
;; NUMBER OF SEQUENCES: 5256  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Human Genome Sciences, Inc.  
;; STREET: 9410 Key West Avenue  
;; CITY: Rockville  
;; STATE: Maryland  
;; COUNTRY: USA  
;; ZIP: 20850  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
;; COMPUTER: HP Vectra 486/33  
;; OPERATING SYSTEM: MSDOS version 6.2  
;; SOFTWARE: ASCII Text  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/329,624  
;; FILING DATE: 27-Dec-2002  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/956,171

```
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 421:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2710 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 421:
US-10-329-624-421

Query Match 60.0%; Score 19.8; DB 17; Length 2710;
Best Local Similarity 75.0%; Pred. No. 3.3e+02;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGGTTGCTTAC 33
Db 558 CCATTAGTAGCATGNTTCTGGCTTGATTAC 589

RESULT 27
US-09-817-182-3
; Sequence 3, Application US/09817182
; Patent No. US20020142431A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001184
; CURRENT APPLICATION NUMBER: US/09/817,182
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 31814
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(31814)
; OTHER INFORMATION: n = A,T,C or G
US-09-817-182-3

Query Match 60.0%; Score 19.8; DB 9; Length 31814;
Best Local Similarity 77.4%; Pred. No. 5.4e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCATTATTACGCTAGTTCCTGGTTGCTT 31
Db 13561 CCATTATCCAGCTACCTCTTGGCTTGCTT 13591

RESULT 28
US-10-473-349-3
; Sequence 3, Application US/10473349
; Publication No. US20040225117A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001184-US
```

```
; CURRENT APPLICATION NUMBER: US/10/473,349
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: PCT/US02/093323
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 09/817,182
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 31814
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(31814)
; OTHER INFORMATION: n = A,T,C or G
US-10-473-349-3

Query Match 60.0%; Score 19.8; DB 18; Length 31814;
Best Local Similarity 77.4%; Pred. No. 5.4e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCATTATTACGCTAGTTCCTGGTTGCTT 31
Db 13561 CCATTATCCAGCTACCTCTTGGCTTGCTT 13591

RESULT 29
US-10-087-192-481
; Sequence 481, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 481
; LENGTH: 39121
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(39121)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-481

Query Match 60.0%; Score 19.8; DB 13; Length 39121;
Best Local Similarity 77.4%; Pred. No. 5.6e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGGTTGCTTAC 33
Db 20031 CATTATGACTCCTGATTCCTGGTTGCTTCC 20061

RESULT 30
US-10-087-192-226
; Sequence 226, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
```

```
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 226
; LENGTH: 653122
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(653122)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-226

Query Match 60.0%; Score 19.8; DB 13; Length 653122;
Best Local Similarity 77.4%; Pred. No. 9.4e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CCATTATTACGCTAGTTCCTGGTTGCTT 31
Db 42506 CACATTATTACTTGGTTTGGTATGCTT 42536

RESULT 31
US-10-723-860-1322
; Sequence 1322, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193 NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1322
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-1322

Query Match 59.4%; Score 19.6; DB 18; Length 310;
Best Local Similarity 84.6%; Pred. No. 2.6e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGGTTT 27
Db 34 CCATTATTACCTAGTCTGGGTTT 59

RESULT 32
US-10-369-493-26447/c
; Sequence 26447, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
```

```
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 26447
; LENGTH: 903
; TYPE: DNA
; ORGANISM: Synechocystis sp.
US-10-369-493-26447

Query Match 58.8%; Score 19.4; DB 17; Length 903;
Best Local Similarity 79.3%; Pred. No. 3.9e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGGTTTGCTT 31
Db 42 CATTATTTTCAGTTGTTCTCTGGTTTGGTT 14

RESULT 33
US-10-367-094-125
; Sequence 125, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 125
; LENGTH: 518360
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(518360)
; OTHER INFORMATION: n = A,T,C or G
US-10-367-094-125

Query Match 58.8%; Score 19.4; DB 18; Length 518360;
Best Local Similarity 79.3%; Pred. No. 1.3e+03;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TTTATTACGCTAGTTCCTGGTTTGCTTAC 33
Db 326381 TTTAAAAAACTAGTTCTCTAGTTTCTTAC 326409

RESULT 34
US-10-461-862-4
; Sequence 4, Application US/10461862
; Publication No. US20050090434A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001800
; CURRENT APPLICATION NUMBER: US/10/461,862
; CURRENT FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 653458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(653458)
; OTHER INFORMATION: n = A,T,C or G
```

```
US-10-461-862-4
Query Match      58.8%; Score 19.4; DB 19; Length 653458;
Best Local Similarity 79.3%; Pred. No. 1.3e+03;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGGTTGCTT 31
|||
Db 158308 CATTATTACGCTTTTGGCTGATGTTT 158336

RESULT 35
US-10-056-052-15
; Sequence 15, Application US/10056052
; Publication No. US20030099656A1
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M
; APPLICANT: HUTCHINS, Jeff T
; APPLICANT: DOMANSKI, Paul
; APPLICANT: PATEL, Pratiksha
; APPLICANT: HALL, Andrea
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN . . .
; FILE REFERENCE: P07069US04/BAS
; CURRENT APPLICATION NUMBER: US/10/056,052
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/308,116
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/298,413
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/274,611
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/264,072
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; TYPE: DNA
; LENGTH: 354
; ORGANISM: Staphylococcus aureus
US-10-056-052-15

Query Match      58.2%; Score 19.2; DB 14; Length 354;
Best Local Similarity 75.0%; Pred. No. 4e+02;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGGTTGCTTAC 33
|||
Db 290 CGCCTACTATGTAAGTTCCTGGTTGCTTAC 321

RESULT 36
US-10-425-115-23753
; Sequence 23753, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 23753
; LENGTH: 411
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(411)
; OTHER INFORMATION: unsure at all n locations
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_121667C.1
US-10-425-115-23753

Query Match      58.2%; Score 19.2; DB 18; Length 411;
Best Local Similarity 75.0%; Pred. No. 4.1e+02;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGGTTGCTTAC 33
|||
Db 27 CCATTTTTCTCAAGTTCCTGGTTGTTC 58

RESULT 37
US-10-161-051-101
; Sequence 101, Application US/10161051
; Publication No. US20030152945A1
; GENERAL INFORMATION:
; APPLICANT: Peter Deak
; APPLICANT: David M Glover
; APPLICANT: Carol Midgley
; TITLE OF INVENTION: Cell cycle progression proteins
; FILE REFERENCE: CCI-021CP
; CURRENT APPLICATION NUMBER: US/10/161,051
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: GB 0007268.6
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-10-161-051-101

Query Match      58.2%; Score 19.2; DB 16; Length 540;
Best Local Similarity 87.5%; Pred. No. 4.3e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 ATTTATTACGCTAGTTCCTGGTTT 27
|||
Db 49 ATTTAGTACGCTCCTTCCTGGTTT 72

RESULT 38
US-10-282-122A-34217/c
; Sequence 34217, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 12:51:52 ; Search time 1685.67 Seconds  
(without alignments)  
745.178 Million cell updates/sec

Title: US-10-089-452-26

Perfect score: 33

Sequence: 1 cccattattacgtagtctctggttgcttac 33

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hcc:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gest1:\*

9: gb\_gest2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	69.7	915	BH135553	ENTOD19TR
2	22	66.7	527	CN076122	EC2BBA11B
3	22	66.7	545	BX745662	BX745662
4	22	66.7	569	AL803878	AL803878
5	22	66.7	571	CN076123	EC2BBA11B
6	22	66.7	597	AL657267	AL657267
7	22	66.7	642	AL659037	AL659037
8	22	66.7	644	AL874808	AL874808
9	22	66.7	732	BX781694	BX781694
10	22	66.7	743	BX690110	BX690110
11	22	66.7	771	AG503427	Mus muscu
12	22	66.7	785	CR585090	CR585090
13	22	66.7	795	CF342840	CF342840
14	22	66.7	851	CF591598	AGENCOURT
15	22	66.7	894	BX688365	BX688365
16	22	66.7	955	BX689719	BX689719
17	21.6	65.5	385	BP072537	BP072537
18	21.4	64.8	566	CB055773	NISC_jj09
19	21.4	64.8	605	DR25G23T	DR25G23T
20	21.4	64.8	628	BB668146	BB668146
21	21.4	64.8	756	BZ611678	WHADS42TF
22	21.4	64.8	3008	AK050663	Mus muscu
23	21	63.6	553	AW642181	cm15b01.w
24	21	63.6	562	AW642281	cm15b10.w

25	21	63.6	571	9	CL596084	CL596084	OB_Ra005
26	21	63.6	616	4	BM119505	BM119505	L0926C02-
27	21	63.6	617	2	AW646111	AW646111	cm61a04.w
28	21	63.6	744	9	CG819443	CG819443	SOYED52TV
29	21	63.6	823	8	BZ487185	BZ487185	BONSU89TF
30	21	63.6	994	7	CO010346	CO010346	EST798681
31	20.8	63.0	237	7	CO90268	CO90268	UNC-pd3ov
32	20.8	63.0	480	8	AQ142791	AQ142791	HS_2230B
33	20.8	63.0	642	8	AQ256169	AQ256169	nbxb0015L
34	20.6	62.4	647	7	CN083910	CN083910	EC2BBA23B
35	20.6	62.4	752	7	CN083909	CN083909	EC2BBA23B
36	20.4	61.8	320	5	BW532470	BW532470	BW532470
37	20.4	61.8	351	1	A1586596	A1586596	486049C07
38	20.4	61.8	485	4	BG267735	BG267735	1000J37C0
39	20.4	61.8	537	6	CD206642	CD206642	HS1_24_E0
40	20.4	61.8	588	7	CO151329	CO151329	EST826382
41	20.4	61.8	607	1	AL890134	AL890134	AL890134
42	20.4	61.8	634	6	CA230495	CA230495	SCJFFL3C0
43	20.4	61.8	638	6	CA120720	CA120720	SCEPLR100
44	20.4	61.8	649	6	CA218458	CA218458	SCSGAD114
45	20.4	61.8	682	6	CD209795	CD209795	HS1_54_A0

## ALIGNMENTS

RESULT 1  
LOCUS BH135553  
DEFINITION ENTOD19TR Entamoeba histolytica Sheared DNA linear GSS 07-AUG-2001  
genomic, genomic survey sequence.  
ACCESSION BH135553  
VERSION BH135553.1 GI:15094614  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica  
ORGANISM Entamoeba histolytica  
Eukaryota; Entamoebidae; Entamoeba.  
REFERENCE 1 (bases 1 to 915)  
Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.  
Determination of clone end sequences from Entamoeba histolytica  
HMI:IMSS sheared DNA library (2001)  
Unpublished (2001)  
JOURNAL Contact: Brendan J Loftus  
COMMENT Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjloftus@tigr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library  
Seq primer: M13-Reverse  
Class: shotgun  
High quality sequence start: 2  
High quality sequence stop: 758.  
Location/Qualifiers  
1..915  
/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HMI:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: pHO51; Site 1: Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark, a  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome

Sequencing: A Practical Approach, eds. M. Vaubin and B. Barell, Oxford University Press, 1999).

ORIGIN  
Query Match 69.7%; Score 23; DB 8; Length 915;  
Best Local Similarity 83.9%; Pred. No. 47;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 3 CATTATTACGCTAGTCTCGTTGCTTAC 33  
|||||  
Db 128 CATTATTATTCTAGTCTCGTTGTTTAC 158  
|||||

RESULT 2  
LOCUS CN076122/c 527 bp mRNA linear EST 30-MAR-2004  
DEFINITION EC2BBA11BF08.b1 Xenopus tropicalis xtb5 plasmid library Xenopus tropicalis cDNA clone xtb511K16 3', mRNA sequence.  
ACCESSION CN076122  
VERSION CN076122.1 GI:45852179  
KEYWORDS EST.  
SOURCE Xenopus tropicalis (western clawed frog)  
ORGANISM Xenopus tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.  
REFERENCE 1 (bases 1 to 527)  
AUTHORS Thuret,R., Fierro,A.C., Coen,L., Perron,M., Demeneix,B., Wegnez,M., Gyapay,G., Weissbach,J., Wincker,P., Mazabraud,A. and Pollet,N.  
TITLE Exploring the nervous system transcriptome in the model Xenopus tropicalis using EST analysis  
JOURNAL Unpublished (2004)  
COMMENT Contact: Pollet N  
Transgenese et Genetique des Amphibiens  
CNRS UMR 8080  
IBATC bat 447, Universite Paris Sud, Orsay, F-91405, France  
Tel: +33 169157272  
Fax: +33 169156816  
Email: Nicolas.Pollet@ibaic.u-psud.fr.  
Location/Qualifiers  
1. 527  
/organism="Xenopus tropicalis"  
/mol\_type="mRNA"  
/strain="ivory coast"  
/db\_xref="taxon:8364"  
/clone="xtb511K16"  
/tissue\_type="pool of brains and spinal cords from tadpoles at stages 51-52 and 61-62"  
/dev\_stage="stage 51-52 and 61-62"  
/lab\_host="E.coli DH10B"  
/clone\_lib="Xenopus tropicalis xtb5 plasmid library"  
/note="Vector: pCMVSPORT6 xtb5; Site\_1: Sfil; Site\_2: Sfil; Xenopus tropicalis polyA+ RNA was obtained from brain and spinal cord of tadpoles at stages 51-52 and 61-62. cDNAs were synthesized using the SMART system of CLONTECH and directionally cloned into pCMVSPORT6 xtb5, a modified version of pCMVSPORT6 allowing directional cloning using asymmetric Sfil sites. For antisense RNA synthesis, use T7 promoter and for sense RNA use SP6 promoter. Library constructed by Dr. L. Coen and Prof. B. Demeneix (Museum National d'Histoire Naturelle and CNRS UMR 5166, Paris, France)."

FEATURES  
source  
ORIGIN  
Query Match 66.7%; Score 22; DB 7; Length 527;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 4 ATTATTACGCTAGTCTCGTTGCTTAC 33  
|||||  
Db 447 ATTATTGCGCTGTTCTCGGTATGTTTC 418  
|||||

ORIGIN  
Query Match 66.7%; Score 22; DB 7; Length 527;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 4 ATTATTACGCTAGTCTCGTTGCTTAC 33  
|||||  
Db 447 ATTATTGCGCTGTTCTCGGTATGTTTC 418  
|||||

RESULT 3  
LOCUS BX745662  
DEFINITION BX745662 XGC-gastrula Xenopus tropicalis cDNA clone TGas054J06 3', mRNA sequence.  
ACCESSION BX745662  
VERSION BX745662.1 GI:38418402  
KEYWORDS EST.  
SOURCE Xenopus tropicalis (western clawed frog)  
ORGANISM Xenopus tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 545)  
AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.  
TITLE Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
JOURNAL Unpublished (2003)  
COMMENT Contact: Croning MDR  
Sanger Institute  
Hinnton, Cambridgeshire, CB10 1SA, UK  
Email: trop@sanger.ac.uk  
Sanger Xenopus tropicalis EST project 2001  
TROPICALIS SEQUENCE ID: TGas054J06.q1kT7  
Sequencing primer: T7  
This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.  
cDNA was oligo dt primed from Sug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCSI07 with Vector: pCSI07; Site 1: EcoRI; Site 2: NotI  
EcoRI at the 5' end and NotI at the 3' end.  
Host: Escherichia coli XL1-blue.  
Location/Qualifiers  
1. 545  
/organism="Xenopus tropicalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8364"  
/clone="TGas054J06"  
/dev\_stage="gastrula (stages 10.5-12 mixed)"  
/lab\_host="Escherichia coli XL1-blue"  
/clone\_lib="XGC-gastrula"  
/note="Vector: pCSI07; Site\_1: EcoRI; Site\_2: NotI; cDNA was oligo dt primed from Sug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCSI07 with EcoRI at the 5' end and NotI at the 3' end."  
end."

FEATURES  
source  
ORIGIN  
Query Match 66.7%; Score 22; DB 5; Length 545;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 4 ATTATTACGCTAGTCTCGTTGCTTAC 33  
|||||  
Db 76 ATTATTGCGCTGTTCTCGGTATGTTTC 105  
|||||

ORIGIN  
Query Match 66.7%; Score 22; DB 5; Length 545;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 4 ATTATTACGCTAGTCTCGTTGCTTAC 33  
|||||  
Db 76 ATTATTGCGCTGTTCTCGGTATGTTTC 105  
|||||

RESULT 4  
LOCUS AL803878  
DEFINITION AL803878 XGC-neurula Xenopus tropicalis cDNA clone TNeu142n06 5', mRNA sequence.  
ACCESSION AL803878  
VERSION AL803878.2 GI:38330021  
KEYWORDS EST.  
SOURCE Xenopus tropicalis (western clawed frog)  
ORGANISM Xenopus tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.  
REFERENCE 1 (bases 1 to 569)  
AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.  
TITLE Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
JOURNAL Unpublished (2003)



```

Query Match      66.7%; Score 22; DB 1; Length 597;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATTTATTACGCTAGTTCCTCGTTGCTTAC 33
Db 126 ATTTATTGCGCTGGTTCCTCGTATGTTTC 155

RESULT 7
AL659037 LOCUS      642 bp mRNA linear EST 07-NOV-2003
DEFINITION AL659037 XGC-neurula Xenopus tropicalis cDNA clone TNeu046d02 5',
mRNA sequence.
ACCESSION AL659037 GI:38226285
VERSION
KEYWORDS
SOURCE
ORGANISM Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL Unpublished (2003)
COMMENT On Dec 13, 2001 this sequence version replaced gi:17672523.

FEATURES
source
1..642
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TNeu046d02"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/clone_lib="XGC-neurula"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."

ORIGIN
Query Match      66.7%; Score 22; DB 1; Length 642;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATTTATTACGCTAGTTCCTCGTTGCTTAC 33
Db 174 ATTTATTGCGCTGGTTCCTCGTATGTTTC 203

RESULT 8
AL874808 LOCUS      644 bp mRNA linear EST 03-DEC-2003
DEFINITION AL874808 XGC-egg Xenopus tropicalis cDNA clone TEGg053l02 5', mRNA
sequence.
ACCESSION AL874808
VERSION
KEYWORDS
EST.

Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 644)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
On Sep 15, 2002 this sequence version replaced gi:22895073.
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
TROPICALIS_SEQUENCE_ID: TEGg053l02.plkasp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from 5ug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XLI-blue.
Location/Qualifiers
1..644
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEGg053l02"
/dev_stage="egg"
/lab_host="Escherichia coli XLI-blue"
/clone_lib="XGC-egg"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

ORIGIN
Query Match      66.7%; Score 22; DB 1; Length 644;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATTTATTACGCTAGTTCCTCGTTGCTTAC 33
Db 385 ATTTATTGCGCTGGTTCCTCGTATGTTTC 414

RESULT 9
BX781694/c LOCUS      732 bp mRNA linear EST 10-DEC-2003
DEFINITION BX781694 XGC-egg Xenopus tropicalis cDNA clone TEGg025n15 3', mRNA
sequence.
ACCESSION BX781694
VERSION
KEYWORDS
SOURCE
ORGANISM Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 732)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
TROPICALIS_SEQUENCE_ID: TEGg025n15.q1kT7
Sequencing primer: T7
This sequence is from a Xenopus Gene Collection (XGC) library

```

constructed by Aaron M. Zorn.  
cDNA was oligo dT primed from Sug of poly A+ RNA from egg.  
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the  
5' end and NotI at the 3' end.  
Vector: pCS107; Site 1: EcoRI; Site 2: NotI  
Host: Escherichia coli XL1-blue.  
Location/Qualifiers

# FEATURES

source

1..732  
/organism="Xenopus tropicalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8364"  
/clone="TEG025n15"  
/dev\_stage="egg"  
/lab\_host="Escherichia coli XL1-blue"  
/clone\_lib="XGC-egg"  
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
was oligo dT primed from Sug of poly A+ RNA from egg.  
EcoRI-NotI cut cDNA was then ligated into pCS107 with  
EcoRI at the 5' end and NotI at the 3' end"

# ORIGIN

Query Match 66.7%; Score 22; DB 5; Length 732;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 ATTATTAGCTAGTTCCTGGTTCTGCTTAC 33  
|||||  
DB 406 ATTATTGCGTGGTTCCTGGTATGTTTC 377

# RESULT 10

BX690110/c

LOCUS

DEFINITION BX690110 XGC-neurula Xenopus tropicalis cDNA clone TNeu088e05 3',  
mRNA sequence.

ACCESSION BX690110

VERSION BX690110.1 GI:38339230

KEYWORDS EST.

SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 743)

REFERENCE Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.

Sanger Xenopus tropicalis EST project 2001 (11\_2003)

Unpublished (2003)

Contact: Croning MDR

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS SEQUENCE ID: TNeu088e05.q1kt7

Sequencing primer: T7

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.

cDNA was oligo dT primed from Sug of poly A+ RNA from neurula.

EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the

5' end and NotI at the 3' end.

Vector: pCS107; Site 1: EcoRI; Site 2: NotI

Host: Escherichia coli DH10B.

# FEATURES

source

1..743  
/organism="Xenopus tropicalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8364"  
/clone="TNeu088e05"  
/dev\_stage="neurula"  
/lab\_host="Escherichia coli DH10B"  
/clone\_lib="XGC-neurula"  
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
was oligo dT primed from Sug of poly A+ RNA from neurula.  
EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end."

# ORIGIN

Query Match 66.7%; Score 22; DB 5; Length 743;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 ATTATTAGCTAGTTCCTGGTTCTGCTTAC 33  
|||||  
DB 491 ATTATTGCGTGGTTCCTGGTATGTTTC 462

# RESULT 11

AG503427/c

LOCUS

DEFINITION

AG503427 Mus musculus molossinus DNA, clone:MSG01-404N18.T7, genomic survey

sequence.

ACCESSION AG503427

VERSION AG503427.1 GI:48210806

KEYWORDS GSS.

SOURCE Mus musculus molossinus

ORGANISM Mus musculus molossinus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
BAC end Sequences of Library MSMg01

1 (bases 1 to 771)

REFERENCE Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.

Unpublished

2 (bases 1 to 771)

Direct Submission

Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suhiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/.

Tel:81-45-503-9111, Fax:81-45-503-9170)

Clones are derived from the mouse BAC library MSMg01. For BAC

library availability, please contact Kuniya Abe (abe@rtc.riken.jp).

Tsukuba Institute, Bio Resource Center,

The Institute of Physical and Chemical Research (RIKEN) 3-1-1

Koyadai, Tsukuba, 305-0074 Japan

phone: 81-298-36-9189, fax: 81-298-36-9199

e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : T7

LIBRARY : pBACe3.6

Vector : EcoRI

R.Site 1 : EcoRI

R.Site 2 : EcoRI.

location/Qualifiers

1..771

/organism="Mus musculus molossinus"

/mol\_type="genomic DNA"

/sub\_species="molossinus"

/db\_xref="taxon:57486"

/clone="MSG01-404N18.T7"

/sex="male"

/tissue type="mixture of kidney and spleen"

/clone\_lib="MSG01 Mouse Male BAC Library"

# ORIGIN

Query Match 66.7%; Score 22; DB 9; Length 771;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CCCATTATTAGCTAGTTCCTGGTTTGTCT 30  
|||||  
DB 111 CTCATTATTAGTACTAGTACTCTGGTTTGTCT 82

# RESULT 12

CR585090/c

LOCUS

CR585090 785 bp mRNA linear EST 26-JUL-2004

```

DEFINITION CR585090 XGC-tailbud-head Xenopus tropicalis cDNA clone THdA046n13
3', mRNA sequence.
ACCESSION CR585090
VERSION CR585090.1 GI:50585090
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 785)
AUTHORS Croning, M.D.R., Ashurst, J.L., Taylor, R., Garrett, N. and Rogers, J.
TITLE Sanger Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
COMMENT Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: THdA046n13.q1k77
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
Seq primer: T7.

FEATURES
Location/Qualifiers
source 1..785
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="THdA046n13"
/dev_stages="tailbud head (stage 28-30)"
/lab_hosts="Escherichia coli DH10B."
/clone_lib="XGC-tailbud-head"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from tailbud
head. EcoRI-NotI cut cDNA was then ligated into pCS107
with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN
Query Match 66.7%; Score 22; DB 7; Length 785;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATTTATTACGCTAGTTCCTGGTTCTGCTTAC 33
|||||
Db 462 ATTTATGCGTGGTTCCTGGTATGTTTC 433

RESULT 13
CF342640/c
LOCUS CF342640
DEFINITION AGENCOURT_15238962 NICHD_XGC_Swb1 Xenopus tropicalis cDNA clone
IMAGE:7000105 3', mRNA sequence.
ACCESSION CF342640
VERSION CF342640.1 GI:33780743
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 795)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Rob Granger, University of Virginia
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

CF342640 795 bp mRNA linear EST 18-AUG-2003
AGENCOURT_15238962 NICHD_XGC_Swb1 Xenopus tropicalis cDNA clone
IMAGE:7000105 3', mRNA sequence.
ACCESSION CF342640
VERSION CF342640.1 GI:33780743
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 795)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Rob Granger, University of Virginia
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

```

```

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14691 row: d column: 24
High quality sequence start: 63
High quality sequence stop: 734.

FEATURES
Location/Qualifiers
source 1..795
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7000105"
/tissue_type="whole body"
/clone_lib="NICHD_XGC_Swb1"
/note="Vector: pExpress-1; Site 1: EcoRV; Site 2: NotI;
Bulk tissue was collected from a whole 10 month old male
from the F6 strain. 1st strand cDNA was primed with a Not
I - oligo(dT) primer, double-stranded cDNA was cloned into
the Not I and EcoRV sites of pExpress-1. Library was
size-selected for >1.5 kb fragments for an average insert
size of 1.92 kb. A normalized version of this library is
also available (NICHD_XGC_SwbIN). Library was constructed
by Open Biosystems (Huntsville, AL). PLEASE NOTE: this
library contains high levels of contamination by worm"

ORIGIN
Query Match 66.7%; Score 22; DB 7; Length 795;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATTTATTACGCTAGTTCCTGGTTCTGCTTAC 33
|||||
Db 455 ATTTATGCGTGGTTCCTGGTATGTTTC 426

RESULT 14
CF591598
LOCUS CF591598
DEFINITION AGENCOURT_15660154 NICHD_XGC_SwbIN Xenopus tropicalis cDNA clone
IMAGE:7023054 5', mRNA sequence.
ACCESSION CF591598
VERSION CF591598.1 GI:36343283
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 851)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Rob Granger, University of Virginia
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14750 row: e column: 04
High quality sequence stop: 688.

FEATURES
Location/Qualifiers
source 1..851
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7023054"
/tissue_type="whole body"

```

/clone lib="NICHG XGC Swbln"  
 /note="Vector: pExpress-1; Site\_1: EcoRV; Site\_2: NotI;  
 Bulk tissue was collected from a whole 10 month old male  
 from the F6 strain. 1st strand cDNA was primed with a Not  
 I - oligo(dT) primer, double-stranded cDNA was cloned into  
 the Not I and EcoRV sites of pExpress-1. Library was  
 size-selected for >1.5 kb fragments for an average insert  
 size of 1.92 kb. Library was normalized to Cot5 with a  
 180-fold reduction of actin. A non-normalized version of  
 this library is also available (NICHG XGC Swbl). Library  
 was constructed by Open Biosystems (Huntsville, AL).  
 PLEASE NOTE: This library contains high level of  
 contamination by worm"

## ORIGIN

Query Match 66.7%; Score 22; DB 7; Length 851;  
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
 Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATTTATTACGCTAGTCTCTGTTGCTTAC 33

Db 560 ATTTATCGCTGGTCTCTGTTGCTTAC 589

RESULT 15  
 BX688365/c  
 LOCUS  
 DEFINITION BX688365 XGC-neurula Xenopus tropicalis cDNA clone TNeu073n20 3',  
 mRNA sequence.

ACCESSION BX688365

VERSION BX688365.1 GI:38337485

KEYWORDS

SOURCE EST.

ORGANISM Xenopus tropicalis (western clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 894)

Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.

Sanger Xenopus tropicalis EST project 2001 (11\_2003)

Unpublished (2003)

Contact: Croning MDR

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS SEQUENCE ID: TNeu073n20.g1kt7

Sequencing primer: T7

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.

cDNA was oligo dt primed from Sug of poly A+ RNA from neurula.

EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the

5' end and NotI at the 3' end.

Vector: pCS107; Site 1: EcoRI; Site\_2: NotI

Host: Escherichia coli DH10B.

Location/Qualifiers

1..894

/organism="Xenopus tropicalis"

/mol\_type="mRNA"

/db\_xref="taxon:8364"

/clone="TNeu073n20"

/dev\_stages="neurula"

/lab\_host="Escherichia coli DH10B"

/clone\_lib="XGC-neurula"

/note="Vector: pCS107; Site 1: EcoRI; Site\_2: NotI; cDNA

was oligo dt primed from Sug of poly A+ RNA from neurula.

EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end."

## ORIGIN

Query Match 66.7%; Score 22; DB 5; Length 894;  
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
 Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATTTATTACGCTAGTCTCTGTTGCTTAC 33

Db 492 ATTTATCGCTGGTCTCTGTTGCTTAC 463

RESULT 16

BX689719/c

LOCUS

DEFINITION BX689719 XGC-neurula Xenopus tropicalis cDNA clone TNeu107c02 3',

mRNA sequence.

ACCESSION BX689719

VERSION BX689719.1 GI:38338839

KEYWORDS

SOURCE EST.

ORGANISM Xenopus tropicalis (western clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;

Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 955)

Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.

Sanger Xenopus tropicalis EST project 2001 (11\_2003)

Unpublished (2003)

Contact: Croning MDR

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS SEQUENCE ID: TNeu107c02.g1kt7

Sequencing primer: T7

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.

cDNA was oligo dt primed from Sug of poly A+ RNA from neurula.

EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the

5' end and NotI at the 3' end.

Vector: pCS107; Site 1: EcoRI; Site\_2: NotI

Host: Escherichia coli DH10B.

Location/Qualifiers

1..955

/organism="Xenopus tropicalis"

/mol\_type="mRNA"

/db\_xref="taxon:8364"

/clone="TNeu107c02"

/dev\_stages="neurula"

/lab\_host="Escherichia coli DH10B"

/clone\_lib="XGC-neurula"

/note="Vector: pCS107; Site 1: EcoRI; Site\_2: NotI; cDNA

was oligo dt primed from Sug of poly A+ RNA from neurula.

EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end."

## ORIGIN

Query Match 66.7%; Score 22; DB 5; Length 955;

Best Local Similarity 83.3%; Pred. No. 1.3e+02;

Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATTTATTACGCTAGTCTCTGTTGCTTAC 33

Db 490 ATTTATCGCTGGTCTCTGTTGCTTAC 461

RESULT 17

BP072537

LOCUS

DEFINITION BP072537 Lotus corniculatus var. japonicus modulating root Lotus

corniculatus var. japonicus cDNA clone GNF070C07 3', mRNA sequence.

ACCESSION BP072537

VERSION BP072537.1 GI:45629248

KEYWORDS

SOURCE EST.

ORGANISM Lotus corniculatus var. japonicus (Lotus japonicus)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

BP072537 385 bp mRNA linear EST 26-AUG-2004  
 BP072537 Lotus corniculatus var. japonicus modulating root Lotus  
 corniculatus var. japonicus cDNA clone GNF070C07 3', mRNA sequence.

BP072537.1 GI:45629248

Lotus corniculatus var. japonicus (Lotus japonicus)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosoids I; Fabales; Fabaceae; Papilionoideae; Loteae;  
 Lotus  
 1 (bases 1 to 385)  
 Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.  
 Characteristics of the Lotus japonicus Gene Repertoire Deduced from  
 Large-Scale Expressed Sequence Tag (EST) Analysis  
 Plant Mol. Biol. 54 (3), 405-414 (2004)  
 Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.  
 Location/Qualifiers  
 1..385  
 /organism="Lotus corniculatus var. japonicus"  
 /mol\_type="mRNA"  
 /isolate="Gifu B-129"  
 /db\_xref="taxon:34305"  
 /clone="GNF070c07"  
 /tissue\_type="nodulating root"  
 /clone\_lib="Lotus corniculatus var. japonicus nodulating  
 root"  
 /note="root of hypernodulating mutant har1-3, rich in  
 nodule primordia"  
 Query Match 65.5%; Score 21.6; DB 5; Length 385;  
 Best Local Similarity 85.7%; Pred. NO. 1.7e+02;  
 Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 Qy 4 ATTATTACGTAGTCTCTGGTTGCTT 31  
 |||||  
 Db 277 ATTATTATCTGGTTCTCTGGTTGTTT 304  
 |||||  
 RESULT 18  
 CB055773 566 bp mRNA linear EST 17-JAN-2003  
 LOCUS NISC-jj09a08.w1 Soares NMBP13-15 Mus musculus cDNA clone  
 DEFINITION IMAGE:4847510 5', mRNA sequence.  
 CB055773  
 VERSION CB055773.1 GI:27794060  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 566)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 cDNA Library Preparation:  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC)  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
 Plate: LLAM10777 row: A column: 15  
 Seq primer: T7 primer.  
 Location/Qualifiers  
 1..566  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4847510"  
 /tissue\_type="pituitary gland"  
 /dev\_stage="juvenile, 13-15 days"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="Soares NMBP13-15"

/note="Organ: brain; Vector: pT7T3D-Paci; Site.1: NotI;  
 Site.2: EcoRI; 1st strand cDNA was primed with a NotI-  
 oligo(dT) primer  
 5'-AATCGGAGATTCGCGCCCTACGATGTTTTTTTTTTT-3';  
 double-stranded cDNA was ligated to EcoRI adaptors  
 5'-AATTCGGCAGCAGG-3' and 5'-CCTCGTCCG-3' (Pharmacia),  
 digested with NotI and cloned into the NotI and EcoRI  
 sites of the pT7T3D-Paci vector. Library went through one  
 round of normalization, and was constructed in the  
 laboratory of M. Bento Soares (University of Iowa)."

## ORIGIN

Query Match 64.8%; Score 21.4; DB 6; Length 566;  
 Best Local Similarity 80.6%; Pred. NO. 2.2e+02;  
 Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 2 CCATTATTACGTAGTCTCTGGTTGCTTA 32  
 |||||  
 Db 325 CTATCCATTAGCTAGTTCAGGTTTGATTA 355  
 |||||

## RESULT 19

DR25G23T 605 bp DNA linear GSS 21-NOV-2002  
 LOCUS Danio rerio genomic clone DKEY-25G23, genomic survey sequence.  
 DEFINITION

ACCESSION AL982728

VERSION AL982728.1 GI:25175048

KEYWORDS GSS.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.

## REFERENCE

AUTHORS Humphray,S.J., Huckle,E. and Hunt,S.E.

TITLE Direct Submission

JOURNAL Submitted (14-NOV-2002)

The Sanger Institute, Wellcome Trust Genome  
 Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:  
[humqurey@sanger.ac.uk](mailto:humqurey@sanger.ac.uk) Unpublished  
 This sequence was generated from the T7 end of BAC 25G23. 25G23 is  
 part of the Daniokey BAC library created by R. Plasterk and N.V.  
 Keygene.

Further details: [http://www.sanger.ac.uk/Projects/D\\_rerio/](http://www.sanger.ac.uk/Projects/D_rerio/).

## FEATURES

source

1..605  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="DKEY-25G23"  
 /tissue\_type="Testis"  
 /note="Vector pIndigoBAC-536"

## ORIGIN

Query Match 64.8%; Score 21.4; DB 9; Length 605;  
 Best Local Similarity 80.6%; Pred. NO. 2.2e+02;  
 Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 3 CATTATTACGTAGTCTCTGGTTGCTTAC 33  
 |||||  
 Db 428 CATTATCAGGCTAATTCGTGATTGCTGAC 458  
 |||||

## RESULT 20

BB668146 628 bp mRNA linear EST 24-OCT-2001  
 LOCUS BB668146 RIKEN full-length enriched, 2 days neonate thymic  
 DEFINITION cells Mus musculus cDNA clone C920030H05.3', mRNA sequence.

ACCESSION BB668146

VERSION BB668146.1 GI:16399595

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;





```

MEDLINE
PUBMED
REFERENCE
AUTHORS
20499374
11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yanamoto,S., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
THE RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
THE FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3008)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohseoto,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
Location/Qualifiers
1..3008
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C920030H05"
/db_xref="taxon:10090"
/clone="C920030H05"
/cell_type="thymus cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
1..3008
/note="zinc finger RNA binding protein (MGD|G1.1341890,

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
20499374
11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yanamoto,S., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
THE RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
THE FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3008)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohseoto,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
Location/Qualifiers
1..3008
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C920030H05"
/db_xref="taxon:10090"
/clone="C920030H05"
/cell_type="thymus cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
1..3008
/note="zinc finger RNA binding protein (MGD|G1.1341890,

FEATURES
source
1..553
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="PBX0114H01"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10B"
/clone_lib="Blackshear/Soares normalized Xenopus egg
library"
/note="vector: pT7T3-Pac; Site 1: EcoRI; Site 2: NotI;
PolyA-selected mRNA was prepared from unfertilized Xenopus
laevis eggs. The library was constructed in the vector
pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and
Soares, M.B. 'Normalization and subtraction: two
approaches to facilitate gene discovery', Genome Research
6:791-806, 1996. The first strand synthesis used a

ORIGIN
GB|NM_011767, evidence: BLASTN, 99%, match=2178)
Query Match 64.8%; Score 21.4; DB 3; Length 3008;
Best Local Similarity 80.6%; Pred. No. 2.7e+02;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 2 CCATTATTACGCTAGTTCCTCGTTGCTTA 32
| | | | | | | | | | | | | | | | | |
Db 2395 CTATCCATTACGCTAGTTCGAGGTTGATTA 2425
| | | | | | | | | | | | | | | | | |

RESULT 23
AW642181
LOCUS AW642181 553 bp mRNA linear EST 26-APR-2001
DEFINITION cm15h01.w1 Blackshear/Soares normalized Xenopus egg library Xenopus
laevis cDNA clone PBX0114H01 5', mRNA sequence.
ACCESSION AW642181
VERSION AW642181.1 GI:7399470
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 553)
AUTHORS Blackshear,P.J., Lai,W.S., Thorn,J.M., Kennington,E.A., Staffa,N.G.
Jr., Moore,D.T., Bouffard,G.G., Beckstrom-Sternberg,S.M.,
Touchman,J.W., Bonaldo,M.F. and Soares,M.B.
The NIEHS Xenopus maternal EST project: interim analysis of the
first 13,879 ESTs from unfertilized eggs
Gene 267 (1), 71-87 (2001)
21211403
11311557
JOURNAL
MEDLINE
PUBMED
COMMENT Contact: Perry J. Blackshear
Office of Clinical Research and Laboratory of Signal Transduction
National Institute of Environmental Health Sciences
A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,
USA
Tel: 919 541-4899
Fax: 919 541-4571
Email: black009niehs.nih.gov
Clone is available through Research Genetics, Inc., 2130 Memorial
Parkway, Huntsville, AL 35901
phone 800-533-4363 ext.cdna, fax 256-536-9016 att:cdna, email
cdna@resgen.com
DNA sequencing and analyses performed by National Institutes of
Health Intramural Sequencing Center (NISC).
PCR Primers
FORWARD: TGTAAACGACGCGCAGT
BACKWARD: CAGGAACAGCTATGACC
Plate: 0114 row: H column: 01
Seq primer: T7 primer.
Location/Qualifiers
1..553
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="PBX0114H01"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10B"
/clone_lib="Blackshear/Soares normalized Xenopus egg
library"
/note="vector: pT7T3-Pac; Site 1: EcoRI; Site 2: NotI;
PolyA-selected mRNA was prepared from unfertilized Xenopus
laevis eggs. The library was constructed in the vector
pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and
Soares, M.B. 'Normalization and subtraction: two
approaches to facilitate gene discovery', Genome Research
6:791-806, 1996. The first strand synthesis used a

```



RESULT 26  
 BML19505/C  
 LOCUS L0926C02-3 NTA Mouse Newborn Kidney cDNA Library (Long) Mus  
 DEFINITION musculus cDNA clone NIA:L0926C02 IMAGE:30002521 3', mRNA sequence.  
 ACCESSION BML19505  
 VERSION BML19505.2 GI:31556383  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 616)  
 Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.  
 TITLE Construction of long-transcript enriched cDNA libraries from  
 submicrogram amounts of total RNAs by a universal PCR amplification  
 method  
 JOURNAL Genome Res. 11 (9), 1553-1558 (2001)  
 MEDLINE 21429098  
 PUBMED 11544199  
 COMMENT On Nov 26, 2001 this sequence version replaced gi:17087531.  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@ngsun.grc.nia.nih.gov  
 Plate: L0926 row: C column: 02  
 Seq primer: -21M13 Forward  
 High quality sequence stop: 616  
 POLYA=Yes.

FEATURES  
 source

1..616  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CS7BL/6J"  
 /db\_xref="niaEST:L0926C02-3"  
 /db\_xref="taxon:10090"  
 /clone="NIA:L0926C02 IMAGE:30002521"  
 /tissue\_type="Newborn Kidney"  
 /dev\_stage="Newborn"  
 /lab\_host="DH10B"  
 /clone\_lib="NTA Mouse Newborn Kidney cDNA Library (Long)"  
 /note="Organ: kidney; Vector: pSPORT1 (Invitrogen);  
 Site 1: SalI; Site 2: NotI; Mouse cDNA project by the  
 Laboratory of Genetics, National Institute on Aging (NIA),  
 Intramural Research Program, NIH  
 (http://lgsun.grc.nia.nih.gov/cDNA). This is a  
 long-transcript enriched cDNA library (Ref. Genome Res.  
 11:1553-1558 (2001). [PMID:11544199]). In  
 brief, double-stranded cDNAs were synthesized with an  
 oligo (dT) primer [Invitrogen: 5'-  
 pGACTAGTCTAGATCGAGCGGCCCTTTT-3'] from 26  
 microgram of total RNA, treated with T4 DNA polymerase,  
 and purified by ethanol-precipitation. The cDNAs were  
 ligated to lone-linker Lb-Sal4, purified by  
 phenol/chloroform, and separated from free linkers by  
 Centricon 100. Then, the cDNAs were amplified by  
 long-range high fidelity PCR using Ex Taq polymerase  
 (Takara) with a primer Sal4-S. The products were purified  
 by phenol/chloroform and Centricon 100. The cDNAs were  
 digested with SalI and NotI enzymes, and cloned into  
 SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.  
 coli host was transformed with ligation mixture by the  
 standard chemical method. The average insert size is about  
 3.0 kb. The library was constructed by Yulan Piao (NIA)."

## ORIGIN

Query Match 63.6%; Score 21; DB 4; Length 616;  
 Best Local Similarity 82.8%; Pred. No. 3.2e+02;  
 Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 4 ATTATTACGCTAGTCTCCTGGTTGCTTA 32

## Db

615 ATCCATTAAAGCTAGTTCACGCTTGATTA 587

## RESULT 27

AW646111

## LOCUS

DEFINITION

AW646111

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

1 (bases 1 to 617)  
 /organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 /clone="PBX0160A04"  
 /sex="female"  
 /tissue\_type="unfertilized egg"  
 /cell\_type="unfertilized egg"  
 /dev\_stage="unfertilized egg"  
 /lab\_host="DH10B"  
 /clone\_lib="Blackshear/Soates normalized Xenopus egg  
 library"  
 /note="Vector: pT7T3-Pac; Site 1: EcoRI; Site 2: NotI;  
 PolyA-selected mRNA was prepared from unfertilized Xenopus  
 laevis eggs. The library was constructed in the vector  
 pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and  
 Soates, M.B. 'Normalization and subtraction: two  
 approaches to facilitate gene discovery'. Genome Research  
 6:791-806, 1996. The first strand synthesis used a  
 NotI-dri8 primer; double stranded cDNAs were ligated to  
 EcoRI adapters, digested with NotI, and directionally  
 cloned into the NotI and EcoRI-digested pT7T3-Pac vector.  
 The library contained approximately 7.2 X 10<sup>5</sup>  
 recombinants, with average insert sizes of 1-1.5 kb."

## ORIGIN

Query Match 63.6%; Score 21; DB 2; Length 617;

[illegible]

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301 838 3519  
Fax: 301 838 0208  
Email: gardner@tigr.org

## FEATURES

source  
Location/Qualifiers  
1..994  
/organism="Coccidioides posadasii"  
/mol\_type="mRNA"  
/strain="C735"  
/db\_xref="taxon:199306"  
/clone="CIEBH07"  
/dev\_stage="spherules"  
/lab\_host="E. coli DH10B, T1 phage resistant"  
/clone\_lib="Coccidioides posadasii spherule cDNA library,  
0.4 to 2.3 kb"  
/note="Vector: pExpress 1; Site 1: Not I; Site 2: Eco RV;  
Coccidioides posadasii spherule cDNA library, 0.4 to 2.3  
kb"

## ORIGIN

Query Match 63.6%; Score 21; DB 7; Length 994;  
Best Local Similarity 82.8%; Pred. No. 3.5e+02;  
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CCCATTATTACGCTAGTTCCTGGTTGC 29  
||||||| ||||||| ||||||| |||||||  
Db 984 CCCATTAGCAGCTAGTTCCTGGTTGC 956

## RESULT 31

CO990268 237 bp mRNA linear EST 09-SEP-2004  
LOCUS UMC-pd3ov2-003-g10 Oviduct gilt D3 of estrous cycle pd3ov Sus  
DEFINITION scrofa cDNA 3', mRNA sequence.  
ACCESSION CO990268  
VERSION CO990268.1 GI:51349542  
KEYWORDS EST.  
SOURCE Sus scrofa (pig)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

## REFERENCE

1 (Bases 1 to 237)  
Kim,J., Whitworth,K.M., Springer,G.K., Bivens,N.J., Ries,J.E.,  
Woods,R.J., Spollen,W.G., Forrester,L.J., Mathialagan,N.,  
Prather,R.S. and Green,J.A.  
Large-scale Generation and Analysis of Expressed Sequence Tags from  
Porcine endometrium and oviduct  
Unpublished (2004)

## JOURNAL

COMMENT Contact: DNA Core Facility (Swine Project)  
Animal Science - RS Prather  
University of Missouri-Columbia  
M616 Medical Sciences Bldg., Columbia, MO 65212, USA  
Tel: (573)882-0428  
Fax: (573)884-5552  
Email: porcine@net.missouri.edu  
POLYA=No.

## FEATURES

source  
Location/Qualifiers  
1..237  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/dev\_stage="Oviduct from a gilt on day 3 of the estrous  
cycle"  
/clone\_lib="pd3ov"

/note="Vector: pSport1; Funding: A grant from the Monsanto  
Company to the University of Missouri. Genetic Source:  
Endometrium and oviduct tissues from various stages of the  
estrous cycle were collected from crossbred pigs (Sus  
scrofa domestica), frozen in liquid nitrogen immediately  
after collection, and stored at -80 degrees Celsius until  
RNA extraction. The specific tissues collected were Day 0  
and Day 3 whole oviducts and Days 3, 6, 10 and 12-14  
endometrium. More information regarding the methods can be

found at:  
<http://genome.rnet.missouri.edu/Swine/Methods/Methods.html>.  
Library Construction (Standard Protocol): All procedures  
discussed in this section have been described in detail  
elsewhere (Soares et al., 1994; Bonaldo et al., 1996;  
Jiang et al., 2001). Total cellular RNA from each sample  
was isolated by using STAT-60 reagent (Tel-Test,  
Friendswood, TX) and poly(A)+ RNA was obtained by two  
rounds of purification with the Oligotex mRNA isolation  
kit (Qiagen) according to the manufacturer's instructions.  
The oviduct libraries and the Day 3, 6 and 10 endometrium  
libraries were constructed essentially as described by the  
manufacturer's instructions provided with the Superscript  
Plasmid System (Invitrogen, cat. no. 18248-013). Briefly,  
1mg of poly(A)+ RNA will be annealed at 37 degrees Celsius  
with 10mg of NotI-tag-dT18 oligonucleotide  
(GTGTCGCGCGCGC-tag-T18) and reverse transcribed at 37  
degrees Celsius with Superscript II (Invitrogen) reverse  
transcriptase (Jiang et al., 2001). The 'tag' represents a  
tissue/stage-specific ten-base sequence identifier  
(<http://genome.uiowa.edu/pubsoft/software.html>) present in  
the oligonucleotide used to prime first-strand synthesis.  
Second strand synthesis was performed with T4 DNA  
polymerase in the presence of DNA ligase and RNase H.  
After second strand synthesis, the double-stranded cDNAs  
was ligated to SalI adapters (Invitrogen-Life  
Technologies) and digested with NotI. The cDNAs will be  
size selected by passage through cDNA size fractionation  
columns (Invitrogen-Life technologies). The cDNAs derived  
from each developmental stage of a particular tissue were  
mixed on an equimolar basis and ligated directionally into  
the NotI and SalI sites of the pSPORT1 vector  
(Invitrogen). After ligation of the inserts, the plasmids  
will be electroporated into DH10B bacteria. The day 12-14  
endometrium library was synthesized by Dr. Bento Soares'  
laboratory (University of Iowa) and was cloned into the  
T37pac vector as described elsewhere (Bonaldo et al.,  
1996). Preliminary Library Characterization: Randomly  
chosen clones from each library were analyzed by  
restriction digestion to determine average insert size (96  
clones) and by sequencing (-4 96-well plates) to confirm  
library quality (e.g. the presence of short polyA+ tails,  
genomic DNA contamination (must be <1%), ribosomal RNA  
clones (must be <1%), etc.) and to provide a sequence  
database representing the predominant clones in each  
library. The clones were sequenced at the University of  
Missouri-Columbia DNA Core Facility. Bioinformatics work  
was performed by GK Springer's bioinformatics group (WG  
Spollen, JE Ries, A Guillen, AA Khambati, RV Patel, CM  
Topinka, SB Bhuiyan) in Computer Science and Health  
Management and Informatics Departments at the University  
of Missouri-Columbia. Clone Requests: Requests for clones  
should be made to the Director of the University of  
Missouri DNA Core facility at: porcine@net.missouri.edu.  
Citations: 1. Bonaldo MF, Lennon G, Soares MB.  
Normalization and Subtraction: Two approaches to  
facilitate gene discovery. Genome Res, 1996; 6:791-806. 2.  
Jiang H, Bivens NJ, Ries JE, Whitworth KM, Green JA,  
Forrester LJ, Springer GK, Didion BA, Mathialagan N,  
Prather RS, Lucy MC (2001) Constructing cDNA libraries  
with fewer clones that contain long poly(dA) tails.  
Biotechniques 31:38-42. 3. Soares MB, MF Bonaldo, P  
Jelene, L Su, L Lawton, A Efstratiadis. 1994.  
Construction and characterization of a normalized cDNA  
library. Proc Natl Acad Sci. 91:9228-9232.  
TAG\_TISSUE=Oviduct gilt D3 of estrous cycle  
TAG\_SEQ=Not found"

## ORIGIN

Query Match 63.0%; Score 20.8; DB 7; Length 237;  
Best Local Similarity 78.1%; Pred. No. 3.5e+02;  
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

Qy 1 CCATTATTACGCTAGTTCCTGGTTGCTTA 32
Db 115 CCATTATTTTCGGTAGTACGGTTGCTGA 146

RESULT 32
AQ142791
LOCUS
DEFINITION HS_2230_B1_G01_MR_CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2230 Col=1 Row=N, genomic survey sequence.
ACCESSION AQ142791
VERSION AQ142791
KEYWORDS GSS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 480)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., and Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,W.D., and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PubMed 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2230 row: N column: 1
Class: BAC ends
High quality sequence stop: 480.
Location/Qualifiers
1..480
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate=2230 Col=1 Row=N"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/Note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

ORIGIN
Query Match 63.0%; Score 20.8; DB 8; Length 480;
Best Local Similarity 78.1%; Pred. No. 3.8e+02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 CCATTATTACGCTAGTTCCTGGTTGCTTAC 33
Db 439 CCATATTCGCTACTTCTCGCTTGCTTAC 470

RESULT 33
AQ256169
LOCUS
DEFINITION nbx0015L17r CUGI Rice BAC Library Oryza sativa (japonica cultivar-group) genomic clone nbx0015L17r, genomic survey sequence.
ACCESSION AQ256169
VERSION AQ256169
KEYWORDS GSS
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
REFERENCE 1 (bases 1 to 642)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
TITLE

REFERENCE
1 (bases 1 to 642)
Wing,R.A. and Dean,R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 5
High quality sequence stop: 191.
Location/Qualifiers
1..642
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="nbx0015L17r"
/tissue type="Leaf"
/lab host="E. coli DH10B"
/clone_lib="CUGI Rice BAC Library"
/Note="Vector: pBeloBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocytledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

ORIGIN
Query Match 63.0%; Score 20.8; DB 8; Length 642;
Best Local Similarity 78.1%; Pred. No. 4e+02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CCCATTATTACGCTAGTTCCTGGTTGCTTA 32
Db 514 CTCATTATTCGCTCTCTCTCTGTTGGTTA 545

RESULT 34
CN083910/c
LOCUS
DEFINITION EC2BBA23BE01.g1 Xenopus tropicalis xtbs plasmid library Xenopus tropicalis CDNA clone xtbs23102 5', mRNA sequence.
ACCESSION CN083910
VERSION CN083910.1
KEYWORDS EST.
SOURCE CN083910.1 GI:45876606
ORGANISM Xenopus tropicalis (western clawed frog)
REFERENCE 1 (bases 1 to 647)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Silurana.
TITLE (Rases 1 to 647)
Thuret,R., Fierro,A.C., Coen,L., Perron,M., Demeneix,B., Wegnez,M., Gvapai,G., Weissenbach,J., Wincker,P., Mazabraud,A. and Pollet,N.
Exploring the nervous system transcriptome in the model Xenopus tropicalis using EST analysis

REFERENCE
1 (bases 1 to 647)
Thuret,R., Fierro,A.C., Coen,L., Perron,M., Demeneix,B., Wegnez,M., Gvapai,G., Weissenbach,J., Wincker,P., Mazabraud,A. and Pollet,N.
Exploring the nervous system transcriptome in the model Xenopus tropicalis using EST analysis

```





```

LOCUS      AI586596      351 bp      mRNA      linear      EST 07-APR-1999
DEFINITION 486049C07.x4 486 - leaf primordia cDNA library from Hake lab Zea
            mays cDNA, mRNA sequence.
ACCESSION  AI586596
VERSION     AI586596.1 GI:4572947
KEYWORDS    EST.
SOURCE      Zea mays
            Zea mays
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 351)
AUTHORS     Walbot,V.
TITLE       Zea mays from various cDNA libraries sequenced at Stanford
JOURNAL     University
COMMENT     Unpublished (1999)
            Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 486049 row: C column: 07.

FEATURES   Location/Qualifiers
            1..351
                /organism="Zea mays"
                /mol_type="mRNA"
                /cultivar="B73"
                /db_xref="taxon:4577"
                /tissue_type="leaf primordia"
                /dev_stage="p7-p11 leaf"
                /lab_host="E.coli XL1-Blue MPR"
                /clone_lib="486 - leaf primordia cDNA library from Hake
                lab"
                /note="Organ: shoot; Vector: Lambda zap; Hake lab cDNA
                library."

ORIGIN
Query Match      61.8%; Score 20.4; DB 1; Length 351;
Best Local Similarity 80.0%; Pred. No. 5.5e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1  CCATTATTACCTAGTCTCGGTTGCT 30
        |||||
        140 CCAGTTTAACTAGTCTAGTTCTCT 169

Db
RESULT 38
BG267735
LOCUS      BG267735      485 bp      mRNA      linear      EST 20-FEB-2001
DEFINITION 1000137C04.x2 1000 - Unigene I from Maize Genome Project Zea mays
            cDNA, mRNA sequence.
ACCESSION  BG267735
VERSION     BG267735.1 GI:12971943
KEYWORDS    EST.
SOURCE      Zea mays
            Zea mays
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 485)
AUTHORS     Walbot,V.
TITLE       Zea mays from various cDNA libraries sequenced at Stanford
JOURNAL     University
COMMENT     Unpublished (1999)
            Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu

```

```

Plate: 1000137 row: C column: 04.
FEATURES   Location/Qualifiers
            1..485
                /organism="Zea mays"
                /mol_type="mRNA"
                /db_xref="dbEST:486049C07.x4"
                /db_xref="taxon:4577"
                /clone_lib="1000 - Unigene I from Maize Genome Project"
            /note="This library represents the unique ESTs found in
            the first round of EST sequencing at Stanford University
            for the maize genome project. Sequences are present from
            libraries 486, 487, 489, 603, 605, 606, 614, 618, 660,
            683, 687, 707, and 945. Contigs were assembled using
            TIGR's CAP program and a representative EST from each
            contig was selected for the Unigene set. All singlets were
            also selected."

ORIGIN
Query Match      61.8%; Score 20.4; DB 4; Length 485;
Best Local Similarity 80.0%; Pred. No. 5.7e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1  CCATTATTACCTAGTCTCGGTTGCT 30
        |||||
        80 CCAGTTTAACTAGTCTAGTTCTCT 109

Db
RESULT 39
CD206642
LOCUS      HSL_24_E04_A012 537 bp      mRNA      linear      EST 20-MAY-2003
DEFINITION clone HSL_24_E04_A012 3', mRNA sequence.
ACCESSION  CD206642
VERSION     CD206642.1 GI:30937913
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
            Sorghum bicolor
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1 (bases 1 to 537)
AUTHORS     Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S.,
            Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Buchanan,C.D.,
            Eastman,A. and Pratt,L.H.
            An EST database from Sorghum: heat-shocked seedlings
            Other ESTs: HSL_24_E04_g1_A012
            Unpublished (2003)
            Contact: Cordonnier-Pratt_MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
            the Human Genome Center, University of Tokyo Institute of Medical
            Science; plant material and RNA prepared at Texas A & M University;
            sequencing done in the Laboratory for Genomics and Bioinformatics,
            University of Georgia. Sequence ends have been trimmed to exclude
            vector and regions below Phred quality 16. Three-prime sequences
            are presented as their reverse complement and have been trimmed to
            exclude polyA.
            Seq primer: Sug3 (CGACCTGCAGCTCGACACA)
            POLYA=No.

FEATURES   Location/Qualifiers
            1..537
                /organism="Sorghum bicolor"
                /mol_type="mRNA"
                /cultivar="IS3620C"
                /db_xref="taxon:4558"
                /clone="HSL_24_E04_A012"
                /lab_host="DH10B-T1 phage-resistant E. coli"
                /clone_lib="Heat-shocked seedlings"

```

/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from 6-day-old seedlings grown in hydroponic culture and heat-shocked at 40-42 C for 4 or 24 hr. After heat shock, roots and leaves were harvested and tissues combined for RNA isolation. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG)."

## ORIGIN

Query Match 61.8%; Score 20.4; DB 6; Length 537;  
 Best Local Similarity 80.0%; Pred. No. 5.8e+02;  
 Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCATTATTACGCTAGTTCCTGGTTGGCTT 31  
 Db 204 CTAATGTTTAATCTAGTTCCTGGTTGGCTT 233

## RESULT 40

CO151329  
 LOCUS  
 DEFINITION  
 EST826382 Aspergillus flavus Normalized cDNA Expression Library  
 Aspergillus flavus cDNA clone NAGEQ51 5' end similar to (Q7SEV8)  
 Predicted protein, mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

CO151329.1 GI:48905330  
 EST. Aspergillus flavus  
 Aspergillus flavus  
 Aspergillus flavus  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
 Yu, J., Whitelaw, C.A., Nierman, W.C., Bhatnagar, D. and Cleveland, T.E.  
 Aspergillus flavus expressed sequence tags for identification of  
 genes with putative roles in aflatoxin contamination of crops  
 FEMS Microbiol. Lett. (2004) In press  
 Contact: Yu J  
 Food and Feed Safety Research Unit  
 USDA/ARS, Southern Regional Research Center  
 1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA  
 Tel: 504 286 4405  
 Fax: 504 286 4419  
 Email: jiu@ars.usda.gov  
 Contact Dr. Yu at USDA/ARS SRRRC (jiu@ars.usda.gov) for clone  
 information  
 PCR Primers  
 FORWARD: M13F  
 BACKWARD: M13R  
 Seq primer: M13 Forward  
 POLYA=No.

## FEATURES

## source

1..588  
 Location/Qualifiers  
 /organism="Aspergillus flavus"  
 /mol\_type="mRNA"  
 /strain="NRRL 3357"  
 /db\_xref="taxon:5059"  
 /clone="NAGEQ51"  
 /sex="asexual mycelia"  
 /cell\_type="mycelia"  
 /dev\_stage="developmental stages from 18 to 96 hours"  
 /lab\_host="E. coli DH10B T1 resistant cells"  
 /clone\_lib="Aspergillus flavus Normalized cDNA Expression Library"  
 /note="Vector: pBlueScript (SK+) (Stratagene), antibiotic selection marker: Carbenicillin; Site 1: NotI, at the 5 prime end; Site 2: EcoRI, at the 3 prime end; This normalized cDNA expression library was constructed using a mixture of mycelial cells grown under eight different medium conditions and harvested at 5 time points (18, 24, 48, 72, 96 hours). The poly-A sequence was trimmed off before ligating to vector."

## ORIGIN

Query Match 61.8%; Score 20.4; DB 7; Length 588;  
 Best Local Similarity 80.0%; Pred. No. 5.8e+02;  
 Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CATTATTACGCTAGTTCCTGGTTGGCTTA 32  
 Db 475 CTTTTTTCCTCTAGTTCCTGGTTTCCTTA 504

Search completed: June 4, 2005, 15:55:52  
 Job time : 1686.67 secs